UNIVERSITY OF CALIFORNIA

Santa Barbara

Mechanisms of rare events in condensed phases

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Chemical Engineering

by

Ryan Gotchy Mullen

Committee in charge:

Professor Baron Peters, Co-Chair

Professor Joan-Emma Shea, Co-Chair

Professor Glenn H. Fredrickson

Professor M. Scott Shell

December 2014

The dissertation of Ryan Gotchy Mullen is approved.

_____

Professor M. Scott Shell

_____

Professor Glenn H. Fredrickson

_____

Professor Joan-Emma Shea, Committee Co-Chair

_____

Professor Baron Peters, Committee Co-Chair

December 2014

Mechanisms of rare events in condensed phases

For MK
Thanks for joining your trajectory to mine.
I don't know where we are going, but I will go there with you.

For Nettie, Rex, Jenna and Marielle
Someday a swarm of loud rambunctious children won't greet me when I return from computerland. That will be a sad day and I'm glad it hasn't happened yet.
Oh, and I attribute half of my gray hair to you. The fun half.

To Joan and Baron
Thank you for collaborating with me on this work. Your questions, insights and guidance have made each project more impactful and each paper more polished.
I have truly gotten the best of both of you by working together.

Curriculum Vita

Ryan Gotchy Mullen
December 2014

Education

B.S. Chemical Engineering, Brigham Young University, Provo, Utah, April 2006 (summa cum laude)
Ph.D. Chemical Engineering, University of California, Santa Barbara, December 2014 (expected)

Professional Employment

Summer 2005: Reservoir Engineering Intern, ExxonMobil
2006-2009: Reservoir Engineer, ExxonMobil
2009: Reservoir Engineer Interim Supervisor, ExxonMobil
2010-2012: Teaching Assistant, Dept. of Chemical Engineering, University of California, Santa Barbara

Awards

NSF Graduate Research Fellowship, University of California, Santa Barbara, 2011
Distinguished Service Award, Dept. of Chemical Engineering, University of California, Santa Barbara, 2011
Teaching Assistant of the Year, Dept. of Chemical Engineering, University of California, Santa Barbara, 2012
Best Oral Presentation (Presentation Skills), Graduate Student Seminar, University of California, Santa Barbara, 2014

Publications

Peters, Baron, Peter G. Bolhuis, Ryan Gotchy Mullen, and Joan-Emma Shea. "Reaction coordinates, one-dimensional Smoluchowski equations, and a test for dynamical self-consistency." *The Journal of chemical physics* 138.5 (2013): 054106.
Mullen, Ryan Gotchy, Joan-Emma Shea, and Baron Peters. "Transmission coefficients, committors, and solvent coordinates in ion-pair dissociation." *Journal of Chemical Theory and Computation* 10.2 (2014): 659-667.
Mullen, Ryan Gotchy, Joan-Emma Shea, and Baron Peters. "Communication: An existence test for dividing surfaces without recrossing." *The Journal of Chemical Physics* 140.4 (2014): 041104.
Mullen, Ryan Gotchy, Joan-Emma Shea, and Baron Peters. "Transition Path Sampling Methods: Flexible Length Aimless & Permutation Shooting" (in preparation)

Presentations

Mullen, Ryan Gotchy, Joan-Emma Shea, and Baron Peters. "Solvent dynamics in ion pair dissociation" Graduate Student Symposium, Santa Barbara, CA (2011)

Mullen, Ryan Gotchy, Joan-Emma Shea, and Baron Peters. "Solvent dynamics in ion pair dissociation" AIChE Annual Meeting, Minnesota, MN (2011)

Mullen, Ryan Gotchy, Joan-Emma Shea, and Baron Peters. "Grote-Hynes, Pollak and dynamics of the committor in ion pair dissociations" Mini-Stat Mech Meeting, Berkeley, CA (2013)

Mullen, Ryan Gotchy, Joan-Emma Shea, and Baron Peters. "Grote-Hynes theory, recrossing and dynamics of the committor" Graduate Student Symposium, Santa Barbara, CA (2013)

Mullen, Ryan Gotchy, Joan-Emma Shea, and Baron Peters. "A simple test for the existence of dividing surfaces without recrossing" AIChE Annual Meeting, San Francisco, CA (2013)

Mullen, Ryan Gotchy, Joan-Emma Shea, and Baron Peters. "An existence test for dividing surfaces without recrossing" American Conference on Theoretical Chemistry, Telluride, CO (2014)

Mullen, Ryan Gotchy, Joan-Emma Shea, and Baron Peters. "Ion pair dissociation: solvent coordinates and reaction rate theories" Graduate Student Symposium, Santa Barbara, CA (2014)

Mullen, Ryan Gotchy, Zach Levine, Gul Zerze, Jeetain Mittal, and Joan-Emma Shea. "Effect of surface hydrophobicity on the conformational stability of GB1 hairpin" AIChE Annual Meeting, Atlanta, GA (2014)

Abstract

Mechanisms of rare events in condensed phases

by

Ryan Gotchy Mullen

Chemical reactions, mass transport in solids, protein folding, and nucleation in first-order phase transitions are examples of processes characterized by multiple, long-lived states. Transitions from one (meta)stable state to another are rare and brief, making them difficult to resolve experimentally. And yet the short transition paths contain valuable structural and dynamic information that governs the lifetimes of the stable states. Transition state theory provides a valuable framework for analyzing rare events, provided that an exact dividing surface with no-recrossing can be found. Direct simulation of rare events processes are complicated by the long waiting times for a transition to spontaneously occur. Simulation methods that introduce a bias decrease the waiting time but also risk altering the mechanism. The reactive flux correlation function provides a two-step recipe for computing rates from simulation using an approximate dividing surface, but may miss important details about the physical reaction mechanism. Transition path sampling (TPS) was developed specifically to sample unbiased dynamical reactive trajectories and in combination with likelihood maximization provides an optimized reaction coordinate model.

We present new, simple TPS methods that reduce the computational expense of simulating rare events over existing methods. We apply these methods to study rare events in

condensed phases and analyze the resulting data with likelihood maximization. The mechanism for vacancy migration in a single domain crystal by activated hops is examined. We find that accurately locating the donor and acceptor sites has a dramatic effect on identifying the mechanism. We next investigate the role of water in ion-pair dissociation, uncovering two solvent mechanisms that influence ion-pair transition states. The resulting dividing surface does not eliminate recrossing, so we present a test for the existence of a no-recrossing surface. It is revealed that an exact dividing surface does not exist for ion-pair dissociation. We discuss the ramifications for transition state theory, Grote-Hynes theory and the relationship between them.

Table of Contents

List of Figures

Chapter 1

Introduction

When the typical transition from one (meta)stable state A to another B is infrequent and much shorter than the lifetime of A or B, the transition is termed a rare event. Rare events are the rate-determining step in chemical reactions [1], molecular isomerizations [2,3] including protein folding [4-8], transport phenomena in solids [9] and glasses [10-12], nucleation in first-order phase transitions [13] and the closely related problem of protein aggregation [5,14]. The short segments connecting A and B, called transition paths, contain valuable structural and dynamic information that governs the lifetimes of the stable states, and yet transition paths cannot be resolved experimentally precisely because they are rare and rapid. Theory and simulation are uniquely able to provide molecular insight into the mechanisms of rare events.

1.1 Transition State Theory

Transitions between A and B are rare because the states are separated by high free energy barriers, as shown in Fig 1.1. Instead of formulating the rate constant k directly from the ensemble of transition paths, transition state theory (TST) [15,16] instead computes the flux through a dividing surface defined by the following characteristics:

1) The dividing surface separates reactants from products.

2) An equilibrium population of states exist on the dividing surface.

3) All trajectories that cross the dividing surface from reactant to product continue on to the product state without ever recrossing.

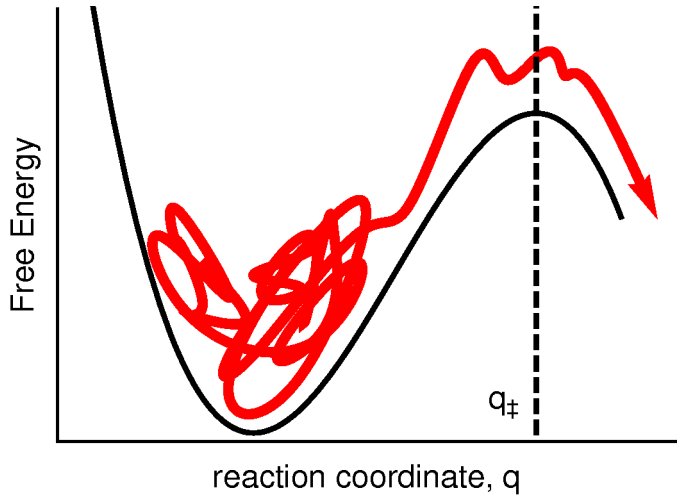

Figure 1.1. Schematic of a free energy barrier projected onto a one-dimensional reaction coordinate q. The dividing surface (dashed) is located at $q_{‡}$. A trajectory (red) that begins in the basin will only rarely and briefly visit the barrier top as it crosses the dividing surface.

When a dividing surface meeting these criteria can be found, the equilibrium one-way flux across the dividing surface, normalized by the population of reactants, is equivalently the non-equilibrium rate constant. For a dividing surface parameterized by $q(\mathbf{x}) = q_{‡}$, where $q(\mathbf{x})$ is the reaction coordinate computed from molecular configuration $\mathbf{x}$ and $q_{‡}$ is the value of $q(\mathbf{x})$ at the dividing surface, the rate constant is

$$k_{\mathrm{TST}}[q] = \frac{\langle \dot{q}\, \theta(\dot{q})\, \delta(q_{‡} - q) \rangle}{\langle \theta(q - q_{‡}) \rangle} \tag{1.1}$$

The brackets indicate an equilibrium average, $\dot{q}$ is the time derivative of q, θ is the Heaviside function and δ is the Dirac delta. Eq. (1.1) reduces to the familiar Arrhenius form

$$k_{\mathrm{TST}}[q] = \frac{1}{2}\langle\,|\dot{q}|\,\rangle_{\ddagger} e^{-\beta\Delta F(q_{\ddagger})} \tag{1.2}$$

where the prefactor is a constrained average of $\dot{q}$ at the dividing surface, $\Delta F(q_{\ddagger})$ is the free energy difference between the ensemble of transition states and reactants, and $\beta = 1/k_{\mathrm{B}}T$. We have emphasized that k_{TST} is a functional of q (and $q_{\ddagger}$). A poor choice of q (or $q_{\ddagger}$) will yield an artificially high k_{TST}, from which Wigner concluded that TST only provides an upper bound to the true rate constant k [17].

In some cases, the dividing surface $q(\mathbf{x}) = q_{\ddagger}$ can be variationally optimized to give the minimum k_{TST} [18,19]. The most successful variational framework is harmonic TST [1], which has been used to describe literally thousands of chemical reactions. In harmonic TST the activation energy and activation entropy are computed from the potential energy surface $U(\mathbf{x})$, often from *ab initio* methods such as density functional theory (DFT). The activation energy is the energy difference between the minimum corresponding to A and the saddle point separating A and B. The activation entropy is computed from the normal mode frequencies at the minimum and at the saddle. The reaction coordinate is the direction corresponding to the one imaginary frequency at the saddle. Harmonic TST is widely used to compute rates, for example, of bond breaking and formation in the gas phase and in surface catalysis [20,21].

For rare events in condensed phases, a dividing surface that is free of recrossing cannot be identified even with harmonic TST [22-24]. The no-recrossing criterion of TST requires that the reaction coordinate dynamics be uncoupled from (or at most weakly coupled to) other modes during the barrier-crossing event. When the reaction coordinate dynamics remain weakly coupled to other modes away from the barrier, as occurs in unimolecular reactions at low pressure, an activated molecule that successfully crosses the barrier may not

thermalize into B before being deflected back across the barrier into A. On the other hand, if the reaction coordinate dynamics are strongly coupled to other modes during the barrier-crossing, as occurs in liquid environments, the reaction coordinate dynamics will exhibit some stochastic character and may therefore recross the dividing surface while still in the transition region. Rates in condensed phases, therefore, often must be computed numerically, for example using molecular dynamics (MD) as we discuss next.

1.2 Simulation of Rare Events

In theory, the rate of A→B transitions could be computed by initializing an MD simulation in A and counting per unit time the number of trajectories that successfully cross over to B. This straight-forward approach, often referred to as 'brute-force' MD, is prohibitively expensive when the energy barrier exceeds a few $k_{\mathrm{B}}T$. A brute-force MD simulation initiated in state A will only sample configurations within a few $k_{\mathrm{B}}T$ of the energy minima. Few, if any, transitions between A and B will be observed.

Instead, the rate constant k can be computed using an arbitrary dividing surface and a transmission coefficient κ that, apart from tunneling corrections, is strictly positive and less than or equal to 1,

$$k = \kappa[q]\, k_{TST}[q] \tag{1.3}$$

In Eq. (1.3) we have emphasized that while κ and k_{TST} are both functionals of q (and $q_{\ddagger}$), their product k is independent of the choice of reaction coordinate or dividing surface, as it must be [25]. Eq. (1.3) suggests a two-part recipe for computing k from MD: first, compute the free energy $F(q)$ and second, compute the prefactor $\kappa[\mathrm{q}]\, \langle\, |\dot{\mathrm{q}}|\, \rangle_{\ddagger} / 2$.

The free energy F projected onto order parameter $q(\mathbf{x})$ is computed by integrating over molecular configurations $\mathbf{x}$

$$e^{-\beta F(q)} = \int d\mathbf{x}\, e^{-\beta U(\mathbf{x})}\, \delta\big(q - q(\mathbf{x})\big) / Z = \big\langle \delta\big(q - q(\mathbf{x})\big)\big\rangle \tag{1.4}$$

where Z is the configurational partition function. As discussed above, configurations corresponding to high free energies will be visited only rarely in brute-force MD. The free energy barrier can be resolved by introducing a potential energy bias along $q(\mathbf{x})$ [26-32]. In this work, we will use umbrella sampling [27] and equilibrium path sampling [33,34] to compute free energy barriers.

κ can be computed, for example, according to Grote-Hynes theory [23] as explained in Chapters 4-5 and Appendix D, or by using the reactive flux correlation function [35]. The latter method is particularly illustrative of the dynamical effect of recrossing. The flux $\dot{q}$ across the dividing surface at time $t = 0$ is correlated with the state of a trajectory at time t,

$$k(t) = \frac{\big\langle \dot{q}(0)\, \theta\big(q(t) - q_\ddagger\big)\, \delta\big(q_\ddagger - q(0)\big)\big\rangle}{\big\langle \theta\big(q - q_\ddagger\big)\big\rangle} \tag{1.5}$$

Note that the reactive flux $k(t)$ of Eq. (1.5) is time-dependent whereas the true reaction rate k of Eq. (1.3) is not. Fig. 1.2 shows the time-dependent behavior of $k(t)$. Initially, $k(0) = k_{\mathrm{TST}}$ because no trajectory has had time to recross the dividing surface. Correlated recrossing of the dividing surface occurs on a timescale t_{mol} that is much shorter than the timescale for independent reactive events t_{rxn}. Accordingly $k(t) \to 0$ as $t \to \infty$ since the current state of a trajectory becomes uncorrelated with $\dot{q}(0)$ at long times. At some intermediate time, the reactive flux plateaus at the value of the true rate constant, $k(t_{\mathrm{plateau}}) = k$, when barrier-crossing trajectories have thermalized into A or B.

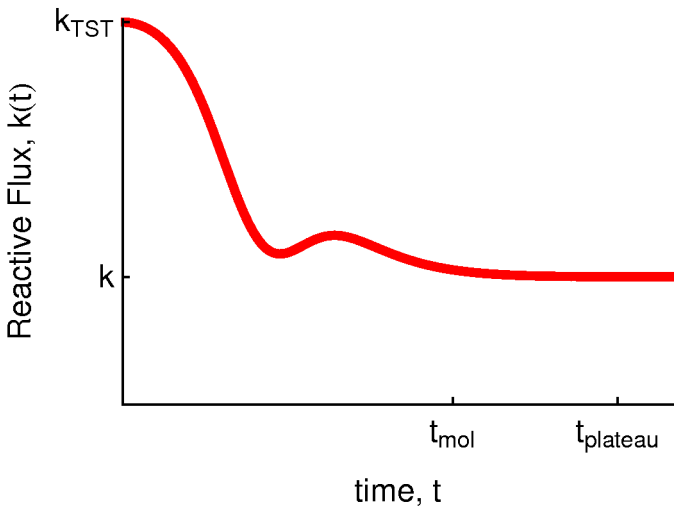

Figure 1.2. Schematic showing the time-dependence of the reactive flux correlation function. The reactive flux initially gives the TST rate constant k_{TST} but decays over the timescale for molecular relaxation t_{mol}. The reactive flux plateaus at the true rate constant k.

κ results from dividing Eq. (1.5) at time $t_{plateau}$ by Eq. (1.1),

$$\kappa[q] = \frac{\langle \dot{q}(0)\, \theta\big(q(t_{plateau}) - q_{\ddagger}\big)\, \delta\big(q_{\ddagger} - q(0)\big)\rangle}{\langle \dot{q}\, \theta(\dot{q})\, \delta\big(q_{\ddagger} - q\big)\rangle} \tag{1.6}$$

The averages in Eq. (1.6) can be computed using short MD simulations initiated from an ensemble of configurations on the putative dividing surface. If $\kappa = 1$, the dividing surface is free of recrossing and $q(\mathbf{x})$ is the reaction coordinate. If $\kappa < 1$, the dividing surface is recrossed, but the source of that recrossing remains unknown. Recrossing may result from the reaction coordinate dynamics being coupled to other modes [22-24], from a poorly chosen dividing surface [17-19], or both. Therefore κ should not be interpreted entirely as a dynamical correction unless the dividing surface has previously been optimized to minimize recrossing.

1.3 Identifying the Reaction Coordinate

Although the reactive flux correlation function allows the rate k to be computed without an exact dividing surface, the statistical accuracy of the method is best if the approximate

dividing surface is close to the exact one. Additionally, since knowing the reaction coordinate is tantamount to knowing the reaction mechanism, an approximate coordinate may obscure important mechanistic details which could aid in manipulating natural processes or designing new ones. Finally, in section 1.1 it was shown that the reaction coordinate q preserves the barrier-crossing dynamics in mapping the high dimensional $U(\mathbf{x})$ to the one-dimensional $F(q)$. A free energy landscape projected onto an approximate coordinate can give inaccurate barriers and distort the microscopic picture of the reaction.

However, identifying a one-dimensional reaction coordinate from among the $3N$ degrees of freedom (in a system with N atoms) often escapes even the keenest intuition. In general, the method is one of trial-and-error. First, a list of trial order parameters is brainstormed and compiled from the literature. Then the best reaction coordinate on the list is found by eliminating poorer coordinates.

A general procedure for evaluating reaction coordinate error was introduced with the development of the committor $p_B(\mathbf{x})$ [36], also know as p_{fold} [37]. The committor $p_B(\mathbf{x})$ is the probability that a trajectory initiated from configuration $\mathbf{x}$ with random momenta will reach state B before reaching A. Transition states $\mathbf{x}_‡$ have equal probability of relaxing to either state, i.e. $p_B(\mathbf{x}_‡) = ½$. Therefore, a histogram of p_B values from states on an exact dividing surface will be infinitely peaked at ½ and zero elsewhere. In practice, committor histograms for good dividing surfaces exhibit some small but finite spread about a mean value $\langle p_B \rangle \approx ½$.

Screening a long list of trial reaction coordinates by computing committor histograms is impractical. First, the free energy $F(q_T)$ must be computed; the dividing surface $q_T(\mathbf{x}) = q_{T‡}$ is typically located at the top of the free energy barrier. Then $p_B(\mathbf{x})$ is computed for a

representative sample of states on the dividing surface by initiating a swarm of trajectories with Boltzmann distributed momenta from each configuration **x** and counting the fraction that relax to B. Computing the $F(q_T)$ and $p_B(\mathbf{x})$ are both expensive procedures. If the resulting histogram is broad, the trial coordinate is a poor one and none of that data can be reused to evaluate other coordinates.

In contrast, likelihood maximization is a method of high-throughput hypothesis testing [38-40]. From a single set of unbiased dynamical trajectories, trial coordinates and linear combinations of trial coordinates can be quickly evaluated and ranked. The best reaction coordinate from likelihood maximization can then be validated by computing the committor histogram, the transmission coefficient, or both.

In the following chapters, we explore mechanisms of barrier-crossing and diffusion-controlled processes. Chapter 2 focuses on the development of new Transition Path Sampling (TPS) methods to reduce the computational cost and complexity of sampling unbiased dynamical trajectories. In Chapter 3, we apply new permutation shooting to vacancy hopping in a crystal lattice. We test two trial reaction coordinates to see which best encapsulates the mechanism at play. In Chapter 4, we investigate the dynamical effect of the solvent on the transmission coefficients and committors of ion-pair dissociation in water. We identify two solvent coordinates that influence the reaction mechanism. Chapter 5 explores the relationship between two rate theories that we use to analyze our ion-pair dissociation results. We present a simple test that uses full multidimensional trajectories to determine which level of theory is needed.

1.4 References

1. Hanggi, P.; Talkner, P.; Borkovec, M. "Reaction-rate theory: fifty years after Kramers." *Rev. Mod. Phys.* **1990**, *62*, 251.
2. Kuharski, R. A.; Chandler, D.; Montgomery Jr, J. A.; Rabii, F.; Singer, S. J. "Stochastic molecular dynamics study of cyclohexane isomerization." *J. Phys. Chem.* **1988**, *92*, 3261-3267.
3. Bolhuis, P. G.; Dellago, C.; Chandler, D. "Reaction coordinates of biomolecular isomerization." *Proc. Natl. Acad. Sci.* **2000**, *97*, 5877-5882.
4. Shea, J.-E.; Brooks III, C. L. "From folding theories to folding proteins: a review and assessment of simulation studies of protein folding and unfolding." *Annu. Rev. Phys. Chem.* **2001**, *52*, 499-535.
5. Dobson, C. M. "Protein folding and misfolding." *Nature* **2003**, *426*, 884-890.
6. Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R. "The protein folding problem." *Ann. Rev. Biophys.* **2008**, *37*, 289.
7. Dill, K. A.; MacCallum, J. L. "The protein-folding problem, 50 years on." *Science* **2012**, *338*, 1042-1046.
8. Englander, S. W.; Mayne, L. "The nature of protein folding pathways." *Proc. Natl. Acad. Sci.* **2014**, *111*, 15873-15880.
9. Nowick, A. S.: *Diffusion in solids: recent developments*; Elsevier, 2012.
10. Angell, C. A.; Ngai, K. L.; McKenna, G. B.; McMillan, P. F.; Martin, S. W. "Relaxation in glassforming liquids and amorphous solids." *J. Appl. Phys.* **2000**, *88*, 3113-3157.
11. Debenedetti, P. G.; Stillinger, F. H. "Supercooled liquids and the glass transition." *Nature* **2001**, *410*, 259-267.
12. Berthier, L.; Biroli, G. "Theoretical perspective on the glass transition and amorphous materials." *Rev. Mod. Phys.* **2011**, *83*, 587.
13. Kashchiev, D.; Van Rosmalen, G. "Review: nucleation in solutions revisited." *Cryst. Res. Technol.* **2003**, *38*, 555-574.
14. Straub, J. E.; Thirumalai, D. "Toward a molecular theory of early and late events in monomer to amyloid fibril formation." *Annu. Rev. Phys. Chem.* **2011**, *62*, 437-463.
15. Eyring, H. "Activated complex in chemical reactions." *J. Chem. Phys.* **1935**, *3*, 107-115.
16. Evans, M. G.; Polanyi, M. "Some applications of the transition state method to the calculation of reaction velocities, especially in solution." *Trans. Faraday Soc.* **1935**, *31*, 875-894.
17. Wigner, E. "Calculation of the Rate of Elementary Association Reactions." *J. Chem. Phys.* **1937**, *5*, 720-725.
18. Keck, J. C. "Variational Theory of Chemical Reaction Rates Applied to Three-Body Recombinations." *J. Chem. Phys.* **1960**, *32*, 1035-1050.
19. Truhlar, D. G.; Garrett, B. C. "Variational transition-state theory." *Acc. Chem. Res.* **1980**, *13*, 440-448.
20. Santen, R. A. v.; Neurock, M.; Shetty, S. G. "Reactivity Theory of Transition-Metal Surfaces: A Brønsted—Evans—Polanyi Linear Activation Energy—Free-Energy Analysis." *Chem. Rev.* **2009**, *110*, 2005-2048.
21. Arnaldsson, A.: Calculation of quantum mechanical rate constants directly from ab initio atomic forces. University of Washington, 2007.
22. Kramers, H. A. "Brownian motion in field of force and diffusion model of chemical reactions." *Physica* **1940**, *7*, 284-304.

23. Grote, R. F.; Hynes, J. T. "The Stable States Picture of Chemical-Reactions .2. Rate Constants for Condensed and Gas-Phase Reaction Models." *J. Chem. Phys.* **1980**, *73*, 2715-2732.
24. Kohen, D.; Tannor, D. J. "Phase-Space Distribution Function Formulation of the Method of Reactive Flux - Memory Friction." *J. Chem. Phys.* **1995**, *103*, 6013-6020.
25. Miller, W. H. "Importance of Nonseparability in Quantum-Mechanical Transition-State Theory." *Acc. Chem. Res.* **1976**, *9*, 306-312.
26. Carter, E.; Ciccotti, G.; Hynes, J. T.; Kapral, R. "Constrained reaction coordinate dynamics for the simulation of rare events." *Chem. Phys. Lett.* **1989**, *156*, 472-477.
27. Torrie, G. M.; Valleau, J. P. "Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling." *J. Comput. Phys.* **1977**, *23*, 187-199.
28. Laio, A.; Parrinello, M. "Escaping free-energy minima." *Proc. Natl. Acad. Sci.* **2002**, *99*, 12562-12566.
29. Maragliano, L.; Vanden-Eijnden, E. "A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations." *Chem. Phys. Lett.* **2006**, *426*, 168-175.
30. Darve, E.; Pohorille, A. "Calculating free energies using average force." *J. Chem. Phys.* **2001**, *115*, 9169-9183.
31. Kong, X.; Brooks III, C. L. "λ-dynamics: A new approach to free energy calculations." *J. Chem. Phys.* **1996**, *105*, 2414-2423.
32. Rosso, L.; Minary, P.; Zhu, Z.; Tuckerman, M. E. "On the use of the adiabatic molecular dynamics technique in the calculation of free energy profiles." *J. Chem. Phys.* **2002**, *116*, 4389-4402.
33. Radhakrishnan, R.; Schlick, T. "Biomolecular free energy profiles by a shooting/umbrella sampling protocol, "BOLAS"." *J. Chem. Phys.* **2004**, *121*, 2436-2444.
34. Peters, B.; Zimmermann, N. E.; Beckham, G. T.; Tester, J. W.; Trout, B. L. "Path sampling calculation of methane diffusivity in natural gas hydrates from a water-vacancy assisted mechanism." *J. Am. Chem. Soc.* **2008**, *130*, 17342-17350.
35. Chandler, D. "Statistical mechanics of isomerization dynamics in liquids and the transition state approximation." *J. Chem. Phys.* **1978**, *68*, 2959-2970.
36. Geissler, P. L.; Dellago, C.; Chandler, D. "Kinetic pathways of ion pair dissociation in water." *J. Phys. Chem. B* **1999**, *103*, 3706-3710.
37. Du, R.; Pande, V. S.; Grosberg, A. Y.; Tanaka, T.; Shakhnovich, E. S. "On the transition coordinate for protein folding." *J. Chem. Phys.* **1998**, *108*, 334-350.
38. Peters, B.; Trout, B. L. "Obtaining reaction coordinates by likelihood maximization." *J. Chem. Phys.* **2006**, *125*, 054108.
39. Peters, B.; Beckham, G. T.; Trout, B. L. "Extensions to the likelihood maximization approach for finding reaction coordinates." *J. Chem. Phys.* **2007**, *127*, 034109.
40. Peters, B. "Inertial likelihood maximization for reaction coordinates with high transmission coefficients." *Chem. Phys. Lett.* **2012**, *554*, 248-253.

Chapter 2

Transition Path Sampling Methods: Flexible Length Aimless & Permutation Shooting

Due to the separation of timescales inherent in rare events, a brute-force molecular dynamics (MD) simulation will include few, if any, transitions between stable states. Some methods overcome this sampling difficulty by introducing a potential energy bias along an *a priori* selected order parameter [1-7]. These methods facilitate the calculation of free energy surfaces and the exploration of configuration space. However, selecting an order parameter that preserves the barrier-crossing dynamics (i.e., the reaction coordinate) is notoriously difficult.

Transition path sampling (TPS) was developed specifically to sample unbiased dynamical reactive trajectories, and therefore enable rate constant calculations without the need for a reaction coordinate [8]. TPS has been used to investigate a wide variety of rare events, for example protein folding [9,10], conformational transitions [11], enzyme catalysis [12,13], nucleation [14-18], amyloid fibril growth [19], micelle budding [20], ligand exchange [21], reactions in solution [22,23], and glassy dynamics [24,25].

The original TPS shooting move [26] generates trial trajectories from an existing transition path using a path based detailed balance criteria. A point on the transition path is randomly selected, the momenta are perturbed, and a trial trajectory is computed from the dynamical equations of motion. The size of the momenta perturbation can be changed to tune the acceptance probability of trial trajectories. As the perturbation decreases, the acceptance increases, with the trade-off that trial trajectories diverge more slowly from the old transition path. However, even with the smallest possible change to the momenta, i.e. at the limit of machine precision, trial trajectories still diverge from the old one within picoseconds [27], making the shooting move "best suited for the study of systems that relax … within the picosecond time scale" [28].

Some methods boost the acceptance probability of longer transition paths by generating trajectories which are similar to the previous trajectory. Examples are precision shooting [28] and one-way shooting [29]. Aimless shooting [30,31] maintains high acceptance and generates rapidly diverging trajectories by a strategy that effectively shoots mostly from the vicinity of the transition states. Peters and Trout [30] showed that a statistical restoring force automatically places the highest diversity of shooting points near the stochastic separatrix, where configurations are most likely to be on a transition path. Since the momenta are drawn afresh from the Boltzmann distribution, aimless shooting is limited to the NVT ensemble and best suited for rare events with stochastic reaction coordinate dynamics.

When barrier-crossing dynamics of the reaction coordinate are uncoupled from other modes, transition paths are best sampled within a microcanonical (NVE) framework. Chemical reactions that form or break covalent bonds have been successfully characterized by harmonic transition state theory, attesting that the bond breaking dynamics are (largely)

independent of other atomic motion. Bond breaking has been simulated in TPS studies of proton hopping [26], ligand exchange [21], methanol coupling [32], and peptide hydrolysis [22]. Even the breaking of a solvated ionic bond, which is part of a diffusional detachment process, has inertial characteristics that are obscured by a stochastic thermostat [23,33]. The permutation shooting move we present in this work rigorously preserves the kinetic energy and linear momentum, making it naturally suited for TPS with NVE dynamics.

Additionally, we formulate aimless shooting and permutation shooting for both fixed-length and for flexible-length TPS. In original TPS, the length of all trajectories is chosen *a priori*. TPS will not give representative transition paths if the chosen length is too long or too short. The shooting move for flexible-length trajectories, introduced by van Erp et al. as part of transition interface sampling [34], eliminated the trajectory length parameter by continuing the time-propagation until a basin is entered. We will demonstrate that flexible-length versions of aimless shooting and permutation shooting are even simpler to implement than for original shooting.

In Section 2.1, we review acceptance criteria that follow detailed balance for original shooting and aimless shooting. In Section 2.2, we develop acceptance criteria for permutation shooting and flexible-length TPS moves.

2.1 Brief Review of Earlier TPS Methods

TPS starts from an initial trajectory that connects states A and B. Each trajectory is an ordered series of N+1 timeslices $\{\mathbf{x}_0, \mathbf{x}_{\Delta t}, \mathbf{x}_{2\Delta t}, \ldots, \mathbf{x}_{N\Delta t}\}$, where $\mathbf{x}$ is a phase space point consisting of configuration $\mathbf{r}$ and momenta $\mathbf{p}$ and the subscript indicates the time along the trajectory. Microscopically, states A and B are defined using one or more order parameters. The order parameter(s) need to distinguish basins A and B and should not overlap. Basin

definitions should be small so that entering trajectories are committed to that thermodynamic state. For the original fixed length TPS algorithms, basins must also be large enough to include typical fluctuations within a stable state, otherwise many trajectories will terminate outside the definitions and be spuriously rejected. A population operator $h_A(\mathbf{x})$ is defined such that for $\mathbf{x}$ in A, $h_A(\mathbf{x}) = 1$, otherwise $h_A(\mathbf{x}) = 0$. Similarly for $h_B(\mathbf{x})$. The population operators in turn are used to define a transition path indicator for trajectory i, $h_{AB}(i) = h_A(\mathbf{x}_0^i)\, h_B(\mathbf{x}_{N\Delta t}^i)$.

Next, a new trajectory n is generated from the initial (old) transition path o. As in Monte Carlo importance sampling of configuration space, n is accepted in TPS according to its statistical weight. As long as n is generated by microscopically reversible dynamics, the probability for accepting n is given by the Metropolis criterion [35]

$$P_{acc}(o \to n) = h_{AB}(n) \min\left[1, \frac{\rho_{eq}(\mathbf{x}_\tau^n)}{\rho_{eq}(\mathbf{x}_\tau^o)} \frac{P_{gen}(n \to \mathbf{x}_\tau^o)}{P_{gen}(o \to \mathbf{x}_\tau^n)}\right] \tag{2.1}$$

where the shooting point is at time τ. In Eq. (2.1), all probabilities related to time propagation along the trajectories have already cancelled as outlined in several earlier works [8,36,37]. Eq. (2.1) only contains the residual factors pertaining to the probability of choosing new and old shooting points from the old and new trajectories, respectively. $\rho_{eq}(\mathbf{x})$ is the equilibrium distribution for phase space point $\mathbf{x}$. In the microcanonical ensemble,

$$\rho_{eq}(\mathbf{x}) = \delta[H(\mathbf{x}) - E]/\,\Omega \tag{2.2}$$

where H is the Hamiltonian, E is the total energy, δ is the Dirac delta function, and Ω is the density of states. In the canonical ensemble, the equilibrium probability is the Boltzmann distribution

$$\rho_{eq}(\mathbf{x}) = \exp\big(-\beta H(\boldsymbol{x})\big)/Z \tag{2.3}$$

where Z is the partition function and β is the inverse temperature $1/k_{\mathrm{B}}T$. $P_{gen}(o \to \mathbf{x}_\tau^n)$ is the probability of generating the new shooting point from the old path and depends on the details of how the timeslice τ is selected on the old trajectory and how the momenta are altered.

2.1.1 Original Shooting

All N+1 timeslices are potential shooting points. The probability of choosing one is $1/N$+1 and is the same for $o \to \mathbf{x}_\tau^n$ and $n \to \mathbf{x}_\tau^o$. The momenta $\mathbf{p}_\tau^o$ are perturbed by adding a random displacement vector $\delta\mathbf{p}$. The components of the displacement vector $\delta\mathbf{p}$ are sampled from a symmetric distribution with the constraint $\sum_i \delta p_i = 0$, such that the probability of choosing $\delta\mathbf{p}$ is equal to that of choosing $-\delta\mathbf{p}$ [26]. The generation probability is symmetric

$$P_{gen}(n \to \mathbf{x}_\tau^o) = P_{gen}(o \to \mathbf{x}_\tau^n) \tag{2.4}$$

Additional constraints, e.g. to preserve angular momentum in systems without periodic boundary conditions or to enforce rigid bonds, require a complicated Gram-Schmidt procedure that, in practice, limits the utility of original shooting.

For microcanonical TPS, the sum $\mathbf{p}_\tau^o + \delta\mathbf{p}$ is rescaled such that $K(\mathbf{p}_\tau^o + \delta\mathbf{p}) = K(\mathbf{p}_\tau^o)$, giving the new momentum $\mathbf{p}_0^n$. From Eqs. (2.1), (2.2) and (2.4), the new trajectory is accepted if it connects basins A and B,

$$P_{acc}(o \to n) = h_{AB}(n) \tag{2.5}$$

For constant temperature TPS, the momenta are not rescaled. Combining Eqs. (2.1), (2.3) and (2.4), the acceptance criteria is

$$P_{acc}(o \to n) = h_{AB}(n) \min\left[1, \exp\left(\beta K(\mathbf{p}_\tau^o) - \beta K(\mathbf{p}_\tau^n)\right)\right] \tag{2.6}$$

where the potential energy Boltzmann factor $\exp(-\beta U(\mathbf{r}_\tau^o))$ is identical for the old and new shooting points.

2.1.2 Aimless Shooting

In aimless shooting, the old trajectory only has two potential shooting points, $\mathbf{r}_0^o$ or $\mathbf{r}_{N\Delta t}^o$, where time $t = 0$ has been shifted to the middle of the trajectory. The probability of choosing a shooting point is ½, for both $o \rightarrow n$ and $n \rightarrow o$. The momenta $\mathbf{p}_\tau^n$ are sampled anew from the Boltzmann distribution. The generation probability is therefore

$$P_{gen}(o \rightarrow \mathbf{x}_\tau^n) = \frac{\exp\left(-\beta K(\boldsymbol{p}_\tau^n)\right)}{2Z} \tag{2.7}$$

which cancels the kinetic energy component of the Boltzmann distribution from Eq. (2.3). From Eqs. (2.1), (2.3) and (2.7), the acceptance criteria for new paths is

$$P_{acc}(o \rightarrow n) = h_{AB}(n) \tag{2.8}$$

2.2 New TPS Methods

2.2.1 Permutation Shooting

Permutation shooting, a new microcanonical TPS algorithm, is easily implemented through the following steps:

1. Randomly select a point on the old trajectory to be the shooting point on the new trajectory $\mathbf{x}_\tau^n$. The potential shooting points on the old trajectory are either $\{\mathbf{x}_\tau^o, \mathbf{x}_{\tau+\Delta t}^o\}$ or $\{\mathbf{x}_\tau^o, \mathbf{x}_{\tau-\Delta t}^o\}$.
2. Randomly select whether the new potential shooting point will be on the backward (at t=-Δt) or forward (at t=Δt) half-trajectory.
3. Select at random two particles with the same mass and swap their momenta, as illustrated in Fig. 2.1. Repeat *i* times.
4. Dynamically propagate the system forward using the new momentum $\mathbf{p}_\tau^n$ and backward in time using the reversed momentum $-\mathbf{p}_\tau^n$.

5. Accept the new trajectory if it joins the reactant and product states, A and B. Else, reject the new trajectory and increase the statistical weight of the old trajectory.

Step 5 reflects that permutation shooting obeys Eqs. (2.2) and (2.4), and therefore has the same acceptance criteria as microcanonical original shooting and canonical aimless shooting

$$P_{acc}(o \to n) = h_{AB}(n) \tag{2.9}$$

However, permutation shooting rigorously preserves the energy E and total momentum and therefore, the corrective procedures in the original shooting move are eliminated. Additionally, the number of momenta swaps i in step 3 is an adjustable parameter that can be tuned to optimize P_{acc}. Randomizing the time step of the alternate shooting point on the new trajectory in step 2 increases the robustness of the algorithm, as discussed in Appendix A.

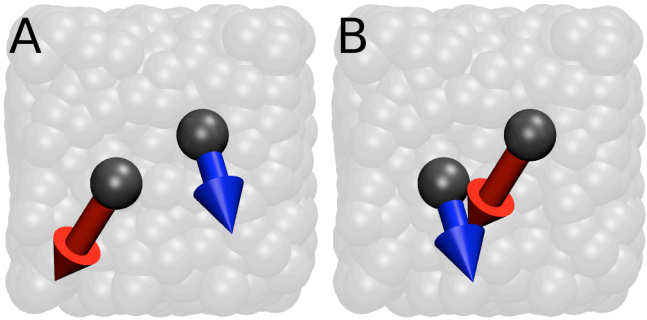

Figure 2.1. Snapshot of an atomic configuration. In permutation shooting, two particles with the same mass are selected at random. **(A)** The original momenta of each particle is indicated by an arrow. **(B)** The momenta have been swapped.

For monoatomic particles and molecules with unconstrained bonds, the atomic momenta can be swapped directly in step 3. Atomic momenta cannot be swapped among molecules with rigid bonds. The constraints that maintain rigid bonds eliminate components of the momenta that would compress or stretch the bond and therefore alter the momenta. In original shooting, the momenta for molecules with rigid bonds are altered using a complicated Gram-Schmidt orthogonalization. In permutation shooting, the solution is much

simpler: swap only the center-of-mass momenta. The center-of-mass momenta are easily computable and none of the components effect bond distances.

2.2.2 Flexible-Length Trajectories

The length of the initial transition path $N\Delta t$ is held constant in original TPS. For long $N\Delta t$, most trajectories needlessly simulate additional time in the stable basins. If $N\Delta t$ is too short, a new trajectory will often terminate, either forward in time or backward in time, without having entered a stable basin. In this case, whether the shooting point $\mathbf{x}_\tau$ is part of a reactive or a non-reactive trajectory remains inconclusive and the transition path ensemble (TPE) will be biased towards shorter trajectories. In a flexible-length trajectory, $\mathbf{x}_\tau$ is propagated in time until the trajectory enters either basin A or B. Each trajectory is just as long as needed to cross the barrier and no longer. Accordingly, the basin definitions should be slightly more restrictive than in fixed-length TPS to ensure the trajectories are committed to the basin upon entry.

Sampling with flexible-length trajectories is especially efficient when the reaction coordinate dynamics are stochastic because some transition paths can be significantly longer than the mean. Juraszek et al. used flexible-length original shooting to study conformational changes of the Trp-cage miniprotein [38]. The TPE for Trp-cage folding exhibits a mean folding time $\langle N\Delta t \rangle$ = 2 ns but also includes 10 ns trajectories. If fixed-length TPS were used, all trajectories would have needed to be 10 ns long in order to accurately sample the TPE.

The method of selecting a shooting point changes the Metropolis criterion for flexible-length TPS. For original shooting, the probability of choosing timeslice τ is $1/(N+1)$ because

every timeslice is a potential shooting point. Since the new and old paths will not be the same length, for microcanonical original shooting with flexible-length trajectories is

$$P_{acc}(o \to n) = h_{AB}(n) \min\left[1, \frac{N^o + 1}{N^n + 1}\right] \tag{2.10}$$

For canonical original shooting with flexible-length trajectories,

$$P_{acc}(o \to n) = h_{AB}(n) \min\left[1, \exp\left(\beta K(\mathbf{p}_\tau^o) - \beta K(\mathbf{p}_\tau^n)\right) \frac{N^o + 1}{N^n + 1}\right] \tag{2.11}$$

For aimless shooting and permutation shooting, the probability of choosing a shooting point is ½ because each trajectory only has two potential shooting points regardless of its length. The acceptance criterion for microcanonical permutation shooting with flexible-length trajectories is

$$P_{acc}(o \to n) = h_{AB}(n) \tag{2.12}$$

For canonical aimless shooting with flexible-length trajectories

$$P_{acc}(o \to n) = h_{AB}(n) \tag{2.13}$$

Thus the acceptance rule for aimless shooting is the same whether the transition paths are flexible-length or fixed-length.

2.3 Conclusions

Aimless shooting and permutation shooting are efficient, simple algorithms to sample the transition paths in the canonical (NVT) and microcanonical (NVE) ensembles, respectively. Permutation shooting rigorously preserves the kinetic energy and total momentum of the shooting point momenta and is simple to implement even for rigid water models. Because each method only uses two potential shooting points per trajectory, the Metropolis acceptance criteria are identical whether or not the sampled transition paths are all the same length.

2.4 References

1. Carter, E.; Ciccotti, G.; Hynes, J. T.; Kapral, R. "Constrained reaction coordinate dynamics for the simulation of rare events." *Chem. Phys. Lett.* **1989**, *156*, 472-477.
2. Torrie, G. M.; Valleau, J. P. "Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling." *J. Comput. Phys.* **1977**, *23*, 187-199.
3. Laio, A.; Parrinello, M. "Escaping free-energy minima." *Proc. Natl. Acad. Sci.* **2002**, *99*, 12562-12566.
4. Maragliano, L.; Vanden-Eijnden, E. "A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations." *Chem. Phys. Lett.* **2006**, *426*, 168-175.
5. Darve, E.; Pohorille, A. "Calculating free energies using average force." *J. Chem. Phys.* **2001**, *115*, 9169-9183.
6. Kong, X.; Brooks III, C. L. "λ-dynamics: A new approach to free energy calculations." *J. Chem. Phys.* **1996**, *105*, 2414-2423.
7. Rosso, L.; Minary, P.; Zhu, Z.; Tuckerman, M. E. "On the use of the adiabatic molecular dynamics technique in the calculation of free energy profiles." *J. Chem. Phys.* **2002**, *116*, 4389-4402.
8. Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. "Transition path sampling: Throwing ropes over rough mountain passes, in the dark." *Annu. Rev. Phys. Chem.* **2002**, *53*, 291-318.
9. Bolhuis, P. G. "Transition-path sampling of ß-hairpin folding." *Proc. Natl. Acad. Sci.* **2003**, *100*, 12129-12134.
10. Juraszek, J.; Bolhuis, P. G. "Rate constant and reaction coordinate of Trp-cage folding in explicit water." *Biophys. J.* **2008**, *95*, 4246-4257.
11. Ma, A.; Dinner, A. R. "Automatic method for identifying reaction coordinates in complex systems." *J. Phys. Chem. B* **2005**, *109*, 6769-6779.
12. Knott, B. C.; Haddad Momeni, M.; Crowley, M. F.; Mackenzie, L. F.; Götz, A. W.; Sandgren, M.; Withers, S. G.; Ståhlberg, J.; Beckham, G. T. "The mechanism of cellulose hydrolysis by a two-step, retaining cellobiohydrolase elucidated by structural and transition path sampling studies." *J. Am. Chem. Soc.* **2013**, *136*, 321-329.
13. Basner, J. E.; Schwartz, S. D. "How enzyme dynamics helps catalyze a reaction in atomic detail: a transition path sampling study." *J. Am. Chem. Soc.* **2005**, *127*, 13822-13831.
14. Moroni, D.; Ten Wolde, P. R.; Bolhuis, P. G. "Interplay between structure and size in a critical crystal nucleus." *Phys. Rev. Lett.* **2005**, *94*, 235703.
15. Beckham, G. T.; Peters, B. "Optimizing Nucleus Size Metrics for Liquid-Solid Nucleation from Transition Paths of Near-Nanosecond Duration." *J. Phys. Chem. Lett.* **2011**, *2*, 1133-1138.
16. Lechner, W.; Dellago, C.; Bolhuis, P. G. "Role of the prestructured surface cloud in crystal nucleation." *Phys. Rev. Lett.* **2011**, *106*, 085701.
17. Pan, A. C.; Chandler, D. "Dynamics of nucleation in the Ising model." *J. Phys. Chem. B* **2004**, *108*, 19681-19686.
18. Jungblut, S.; Dellago, C. "Crystallization of a binary Lennard-Jones mixture." *J. Chem. Phys.* **2011**, *134*, 104501.
19. Schor, M.; Vreede, J.; Bolhuis, P. G. "Elucidating the locking mechanism of peptides onto growing amyloid fibrils through transition path sampling." *Biophys. J.* **2012**, *103*, 1296-1304.

20. Pool, R.; Bolhuis, P. G. "Sampling the kinetic pathways of a micelle fusion and fission transition." *J. Chem. Phys.* **2007**, *126*, 244703.
21. Snee, P. T.; Shanoski, J.; Harris, C. B. "Mechanism of ligand exchange studied using transition path sampling." *J. Am. Chem. Soc.* **2005**, *127*, 1286-1290.
22. Pan, B.; Ricci, M. S.; Trout, B. L. "A molecular mechanism of hydrolysis of peptide bonds at neutral pH using a model compound." *J. Phys. Chem. B* **2011**, *115*, 5958-5970.
23. Mullen, R. G.; Shea, J.-E.; Peters, B. "Transmission coefficients, committors, and solvent coordinates in ion-pair dissociation." *J. Chem. Theory Comput.* **2014**, *10*, 659-667.
24. Keys, A. S.; Hedges, L. O.; Garrahan, J. P.; Glotzer, S. C.; Chandler, D. "Excitations Are Localized and Relaxation Is Hierarchical in Glass-Forming Liquids." *Phys. Rev. X* **2011**, *1*.
25. Xi, L.; Shah, M.; Trout, B. L. "Hopping of water in a glassy polymer studied via transition path sampling and likelihood maximization." *J. Phys. Chem. B* **2013**, *117*, 3634-3647.
26. Geissler, P. L.; Dellago, C.; Chandler, D. "Chemical dynamics of the protonated water trimer analyzed by transition path sampling." *Phys. Chem. Chem. Phys.* **1999**, *1*, 1317-1322.
27. Allen, R. J.; Frenkel, D.; ten Wolde, P. R. "Simulating rare events in equilibrium or nonequilibrium stochastic systems." *J. Chem. Phys.* **2006**, *124*, -.
28. Grünwald, M.; Dellago, C.; Geissler, P. L. "Precision shooting: Sampling long transition pathways." *J. Chem. Phys.* **2008**, *129*, 194101.
29. Bolhuis, P. G. "Transition path sampling on diffusive barriers." *J. Phys.: Condens. Matter* **2003**, *15*, S113.
30. Peters, B.; Trout, B. L. "Obtaining reaction coordinates by likelihood maximization." *J. Chem. Phys.* **2006**, *125*, 054108.
31. Peters, B.; Beckham, G. T.; Trout, B. L. "Extensions to the likelihood maximization approach for finding reaction coordinates." *J. Chem. Phys.* **2007**, *127*, 034109.
32. Lo, C. S.; Radhakrishnan, R.; Trout, B. L. "Application of transition path sampling methods in catalysis: A new mechanism for C-C bond formation in the methanol coupling reaction in chabazite." *Catal. Today* **2005**, *105*, 93-105.
33. Ballard, A. J.; Dellago, C. "Toward the Mechanism of Ionic Dissociation in Water." *J. Phys. Chem. B* **2012**, *116*, 13490-13497.
34. van Erp, T. S.; Moroni, D.; Bolhuis, P. G. "A novel path sampling method for the calculation of rate constants." *J. Chem. Phys.* **2003**, *118*, 7762-7774.
35. Dellago, C.; Bolhuis, P. G.; Chandler, D. "On the calculation of reaction rate constants in the transition path ensemble." *J. Chem. Phys.* **1999**, *110*, 6617-6625.
36. Dellago, C.; Bolhuis, P. G.; Geissler, P. L.: Transition path sampling methods. In *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1*; Springer, 2006; pp 349-391.
37. Dellago, C.; Bolhuis, P. G.: Transition path sampling and other advanced simulation techniques for rare events. In *Advanced computer simulation approaches for soft matter sciences III*; Springer, 2009; pp 167-233.
38. Juraszek, J.; Vreede, J.; Bolhuis, P. G. "Transition path sampling of protein conformational changes." *Chem. Phys.* **2012**, *396*, 30-44.

Chapter 3

Vacancy Diffusion in a Two-dimensional Lennard-Jones Crystal

Diffusion in single crystal domains occurs as atoms jump through interstitial sites, vacancies, or a combination of the two. Vacancy assisted diffusion is important for dopant infusion in silicon [1-3], mass transfer in gas hydrates [4-6], growth of nanostructured materials [7,8], oxygen migration in TiO_2 catalysts [9], radiation damage and creep in steels [10,11], and many other processes.

At temperatures approaching the melting point, vacancy diffusion is a complex process that may involve chains of correlated jump events, vacancy-vacancy interactions, or collective lattice reorientation. At low temperatures, vacancy diffusion is a rare event occurring only when a nearest neighbor particle hops into the vacant lattice site. It this chapter, we restrict our work to the low temperature regime.

In section 3.1, we describe the model crystal we employ. In section 3.2, we present the two order parameters that we will test. In section 3.3, we implement transition path sampling (TPS) using permutation shooting. Then we present our results (section 3.4) and draw conclusions (section 3.5).

3.1 Model

We applied flexible-length permutation shooting to study vacancy migration in a two-dimensional trigonal crystal of Lennard-Jones (LJ) particles. Our system consists of 15 x 10 trigonal unit cells populated with 299 identical LJ particles of mass m, leaving one vacant lattice site. Neighboring lattice sites are spaced $2^{1/6}\sigma$ apart, which corresponds to the minimum energy $-\epsilon$ of the Lennard-Jones potential. Dynamics were simulated using LAMMPS [12] with a 0.001 τ_{LJ} timestep, where $\tau_{LJ} = (m\sigma^2/\epsilon)^{1/2}$, and periodic boundary conditions.

A vacancy is surrounded by 6 particles, which are each equally likely to move into the vacancy. We removed this 6-fold degeneracy by tagging one particle H and studying only trajectories or configurations for which H is the hopping particle. Before a hopping event, H is located at a donor site D and is surrounded by five neighboring particles, which we numbered 1-5. After a hopping event, H is at an acceptor site A and still has five neighbors, particles 4 and 5 plus three other particles we numbered 6-8 (see Fig. 3.1).

To locate the vacancy before and after a jump, we used a scheme similar to Geslin et al. [13] (see also Ref. [6]). A Weeks-Chandler-Anderson (WCA) [14] probe particle with the same σ and ϵ that governs the LJ interactions was used to sample the potential field generated by each crystal configuration. This particle is sufficiently large that the global minima located at the unoccupied lattice site is easily distinguishable from the local minima between occupied lattice sites. At low densities or in a crystal of hard-sphere particles, the two different minima may be similar in depth. Geslin et al. locate the vacancy even in these cases

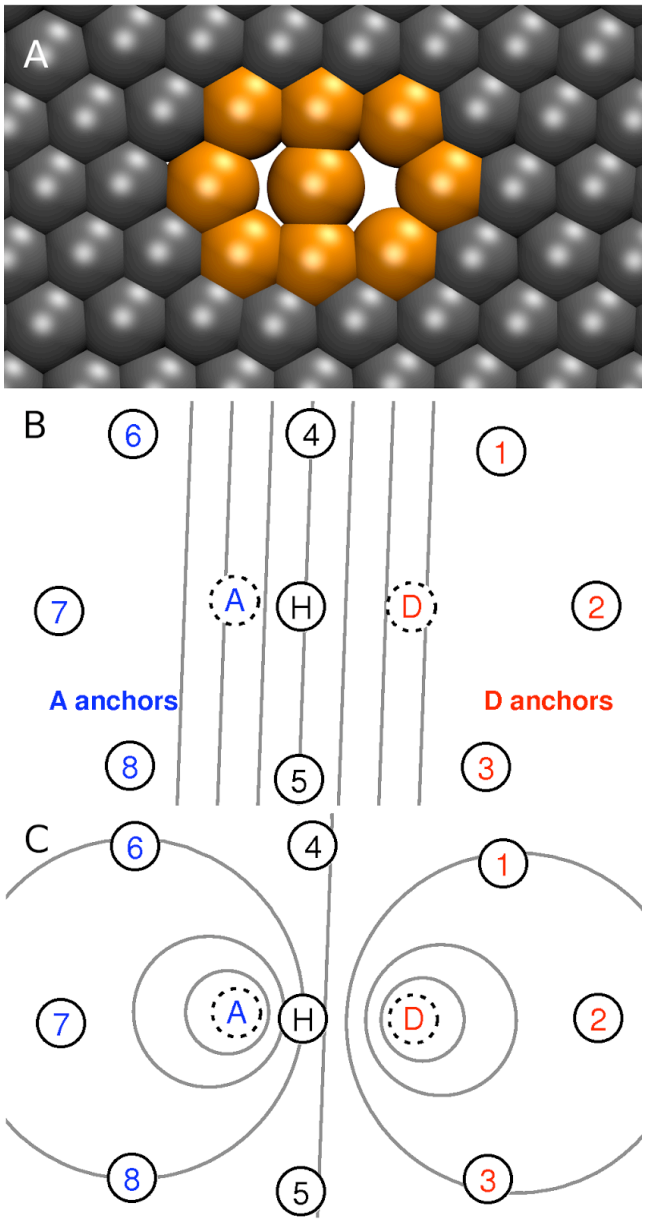

Figure 3.1. (A) Snapshot of a configuration near the transition state. The hopping particle and its eight neighbors are highlighted orange. **(B)** Diagram showing the highlighted particles, donor site *D*, acceptor site *A* and isosurfaces of order parameter ξ (see text). Particle *H* lies approximately on isosurface $\xi = 0$ and isosurfaces are shown at intervals of 0.25. **(C)** Same configuration of particles with isosurfaces of coordinate q (see text). *H* lies on isosurface $q = 0.5$ and isosurfaces of q are shown in intervals of 0.5.

by transforming the classical potential field into a pseudo-quantum ground state density. For the present case, it is sufficient to locate the density at the global minima of the classical potential field. We enforce the restriction that *H* is the only hopping particle by rejecting any

configurations for which the vacancy is surrounded by particles other than {*H*,4-8} or {*H*,1-5}.

To track the progress of the hopping event, we need to locate *D* and *A*, even when *H* is at neither one. In our first attempt, we define *D* as the minimum energy position of *H* considering only the interactions of *H* with particles 1-3 (the "*D* anchors") fixed in their current positions. *A* is similarly defined with respect to particles 6-8, the "*A* anchors" (see Fig. 3.1B). A hopping event therefore begins with *H* fluctuating around *D* and ends with *H* fluctuating around *A*. Our results will show that this energy minimization method for locating *D* and *A* is not sufficiently accurate for computing trial reaction coordinate velocities. However, it does distinguish pre-hop, hopping, and post-hop configurations, which is the only way we use it here.

3.2 Trial Order Parameters

We evaluate two order parameters that have been used to describe vacancy diffusion. The first order parameter we consider was formulated by Bennett to study vacancy diffusion in a face-centered cubic (FCC) crystal of LJ particles. [15] Bennett recognized that the hopping particle must pass through a gate composed of other nearest neighbor particles. He proposed an order parameter that tracks the position of the hopping particle relative to the center of mass of the gate particles

$$\xi = \left(\mathbf{r}_H - \frac{1}{N_g} \sum_i \mathbf{r}_i \right) \cdot \hat{\mathbf{r}}_{AD} \tag{3.1}$$

where $\mathbf{r}_H$ is the position of the hopping particle, the N_g gate particles are at positions $\mathbf{r}_i$, and $\hat{\mathbf{r}}_{AD}$ is the unit vector in the direction from the D to A. The $\xi = 0$ isosurface is the putative dividing surface separating pre-hop ($\xi < 0$) and post-hop ($\xi > 0$) configurations.

The second order parameter was used by Peters et al. to study methane hopping between clathrate cages in hydrates [6]. The bipolar coordinates

$$q = \ln\left(\frac{r_{HD}}{r_{HA}}\right) \tag{3.2}$$

where r_{HD} and r_{HA} are the distance from H to D and from H to A, respectively, and the angle

$$\theta = \angle DHA \tag{3.3}$$

describe the position of the particle hopping between donor and acceptor cages. The $q = 0$ and $\xi = 0$ isosurfaces shown in Fig. 3.1B and 3.1C are nearly, but not exactly, identical. The $\xi = 0$ isosurface depends on A, D and separately on the gate atom positions. The $q = 0$ isosurface depends only on the A and D positions.

3.3 Methods

We performed N_R = 6,000 permutation shooting moves at an energy E = -826.5ϵ. The two candidate shooting points were separated by $\Delta t = 25$ timesteps. A new shooting point momenta $\mathbf{p}^n$ was generated by swapping $k = 299$ momenta pairs between particles. Based on previous applications of q, the symmetry of q and on typical unbiased values of q, we defined the donor basin as $q < -3$ and the acceptor basin is $q > 3$. The transition path ensemble consists of 1638 unique trajectories with an average temperature $\langle T \rangle = 0.199\ \epsilon/k_B$ and an average trajectory length $\langle N\Delta t \rangle = 1.26\ \tau_{LJ}$ with a standard deviation of 0.30 τ_{LJ}.

The order parameters ξ and q were tested by inertial likelihood maximization [16] (iLMax) to determine the best reaction coordinate. iLMax optimizes trial coordinates to

predict the committor p_B and give a high transmission coefficient κ. p_B is the probability that a trajectory initiated from a configuration with Boltzmann distributed momenta will end in the product state. A high κ, approaching the classical upper limit $\kappa = 1$, indicates that recrossing of the transition state dividing surface has been effectively eliminated.

Trial coordinates have the form $q_T(z) = c_0 + c_1 z$, where z is either ξ or q and coefficients c_0 and c_1 are adjustable parameters. The trial coordinate is added to the velocity $\dot{q}_T(z)$ in the reaction probability

$$p_{RX} = \frac{1 + \mathrm{erf}[q_T(z) + c_V\, \dot{q}_T(z)]}{2} \tag{3.4}$$

z is computed for each shooting point configuration ($t = 0$). The velocities $\dot{z}$ are computed by finite difference using the shooting point configuration and the configuration at $t = 1$ timestep. The coefficients c_0, c_1 and c_V are adjusted to maximize the likelihood L

$$L = \prod_{\mathbf{x}\in \mathrm{B}} p_{RX} \prod_{\mathbf{x}\in \mathrm{A}} (1 - p_{RX}) \tag{3.5}$$

where the first product is over shooting points that end in B and the second product is over points that end in A. The best reaction coordinate is indicated by the highest L. Adding a variable to the trial coordinate is only significant if the log-likelihood increases by several increments of a Bayesian criterion $\delta L_{\min} = \frac{1}{2} \ln(N_R)$, where N_R is the number of trajectories ($N_R = 6{,}000$ in this work).

The free energy was computed for $q \geq 0$ by equilibrium path sampling (EPS [6], also known as BOLAS [17]) at temperature $T = 0.2\, \epsilon/k_B$. Temperature was controlled by a Nose-Hoover thermostat [18,19] using a damping time of 10 timesteps. Since q is symmetric, the free energy $q \leq 0$ will be identical. Along the barrier, $0.0 \leq q \leq 2.3$, we used 9 EPS windows each $q_{i,max}$-$q_{i,min} = 0.30$ wide and overlapping by $q_{i+1,min}$-$q_{i,max} = 0.05$. We harvested 50,000

trajectories within each window. For the window at the barrier top, each trajectory was 200 timesteps long. On the downward slope of the barrier, $0.25 \leq q \leq 2.25$, trajectories were shortened to 100 timesteps. The free energy in the acceptor basin was calculated using a 10th EPS window, $2.25 \leq q \leq 3.80$, using 100,000 trajectories 1000 timesteps long. In every case, 11 equally spaced configurations were saved from each trajectory. The resulting free energy $F(q)$ was shifted vertically so that the free energy of the reactants $\int_{-6}^{0} dq\, e^{-\beta F(q)} = 1$.

Likelihood maximization results were validated by computing $\kappa[\xi]$ and $\kappa[q]$ from the reactive flux correlation function. An ensemble of 1,000 configurations from the $\xi = 0$ and $q = 0$ dividing surfaces were randomly selected from the EPS trajectories at the barrier top. Momenta were drawn from the Boltzmann distribution at $T = 0.2\ \epsilon/k_B$. Forward and back trajectory pairs were computed for 1 τ_{LJ}.

3.4 Results

Table 3.1 shows the log-likelihood scores for selected trial coordinates. Log-likelihood scores are reported relative to the value of $\ln L[q]$ in increments of δL_{min}, i.e. $\Delta\ln L[q_T]=(\ln L[q_T]+2129)/4.35$. The preliminary score $\Delta\ln L[\xi] = 178$ shows that ξ is better than q as a reaction coordinate.

To understand why ξ is better than q, we compare Eqs. (3.1) and (3.2). Both ξ and q depend on the coordinate $\mathbf{r}_H$. q is also highly sensitive to the location of D and A, whereas ξ only depends on the direction from D to A. We compute velocities from configurations one timestep apart. The displacement of H over one timestep, averaged over the $N_R = 6000$ shooting points, is $\langle|\Delta\mathbf{r}_H|\rangle = 6.4 * 10^{-4}\sigma$. By contrast, the average displacements of D and A are nearly an order of magnitude larger, $\langle|\Delta\mathbf{r}_D|\rangle = 1.5 * 10^{-3}\sigma$ and $\langle|\Delta\mathbf{r}_A|\rangle = 1.9 * 10^{-3}\sigma$. Large

fluctuations in $\mathbf{r}_D$ and $\mathbf{r}_A$ cause large fluctuations in $\dot{q}$ which in turn obscure any information which would otherwise by conveyed by the velocity. One plausible explanation for the large fluctuations in $\mathbf{r}_D$ and $\mathbf{r}_A$ is that the potential energy used to locate *D* and *A* could have several minima nearly degenerate in energy, such that a small perturbation of particles 1, 2, 3, 6, 7, or 8 leads to a large shift in the global minimum.

Table 3.1. Inertial Likelihood Maximization Results

Reaction Coordinate	c_V	$\Delta \ln L$
***D* & *A* located by energy minimization**		
$0.11 + 8.4\,\xi$	0.16	178
$0.026 + 1.2\,q$	0.040	0[a]
***D* & *A* located by linear combination of surrounding atoms**		
$0.14 + 11\,\xi_{LC}$	0.18	231
$0.064 + 4.0\,q_{LC}$	0.17	310

The best reaction coordinate is indicated by the highest log-likelihood, $\Delta \ln L$ (see text).
[a] $\ln L[q] = -2129$

As an alternative, we locate *D* and *A* using lattice vectors from the positions of nearby particles. Any given particle fluctuates randomly around its lattice site, so the accuracy of estimating $\mathbf{r}_A$ and $\mathbf{r}_D$ will increase with the number of reference particles. We compute the lattice sites *D* and *A* from the current positions of the six anchor particles,

$$\mathbf{r}_D = (3\,\mathbf{r}_1 + 2\,\mathbf{r}_2 + 3\,\mathbf{r}_3 + \mathbf{r}_6 + \mathbf{r}_7 + \mathbf{r}_8) \,/\, 11 \tag{3.6}$$

and

$$\mathbf{r}_A = (\mathbf{r}_1 + \mathbf{r}_2 + \mathbf{r}_3 + 3\,\mathbf{r}_6 + 2\,\mathbf{r}_7 + 3\,\mathbf{r}_8) \,/\, 11 \tag{3.7}$$

Using Eqs. (3.6) and (3.7), the average displacements $\langle|\Delta\mathbf{r}_D|\rangle = 2.5 * 10^{-4}\sigma$ and $\langle|\Delta\mathbf{r}_A|\rangle = 2.5 * 10^{-4}\sigma$ are comparable to $\langle|\Delta\mathbf{r}_H|\rangle$. Trial coordinates computed using Eqs. (3.6) and (3.7) have the subscript 'LC'. iLMax scores for ξ_{LC} and q_{LC} are both improvements over ξ and q, respectively, with the final result that q_{LC} is the top coordinate.

With the optimized reaction coordinate q_{LC}, it is now straight-forward to compute the rate

$$k = \kappa[q_{LC}]\frac{\langle|\dot{q}_{LC}|\rangle_{\ddagger}}{2}e^{-\beta\Delta F\left(q_{LC_{\ddagger}}\right)} \tag{3.8}$$

where $\kappa[q_{LC}]$ is the transmission coefficient, $\langle|\dot{q}_{LC}|\rangle_{\ddagger}$ is the equilibrium average velocity of q_{LC} constrained to the dividing surface $q_{LC}(\mathbf{x}) = q_{LC\ddagger}$, and $\Delta F(q_{LC\ddagger})$ is the free energy difference between the ensemble of transition states and the ensemble of reactants. The free energy $F(q_{LC})$ (not shown) was estimated by reweighting the EPS data and exhibits only small deviations from $F(q)$. $F(q)$ shows a barrier at $q = 0$ that is $\Delta F(q_{\ddagger}) = 11.0\ k_B T$ above the free energy of the reactant state (see Fig. 3.2). The prefactor $\kappa[q]\langle|\dot{q}|\rangle_{\ddagger} / 2$ was computed by the method of reactive flux [20] to be $0.660 / \tau_{LJ}$. The rate constant is $k = 1.10 * 10^{-5} / \tau_{LJ}$.

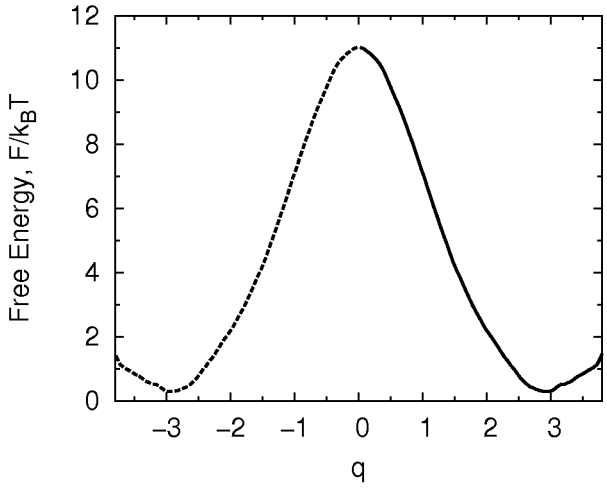

Figure 3.2. The free energy projected onto coordinate q. The dotted portion is the symmetric image of the computed $q > 0$ curve (solid). The integrated free energy of the reactant state $q \leq 0$ was set to zero.

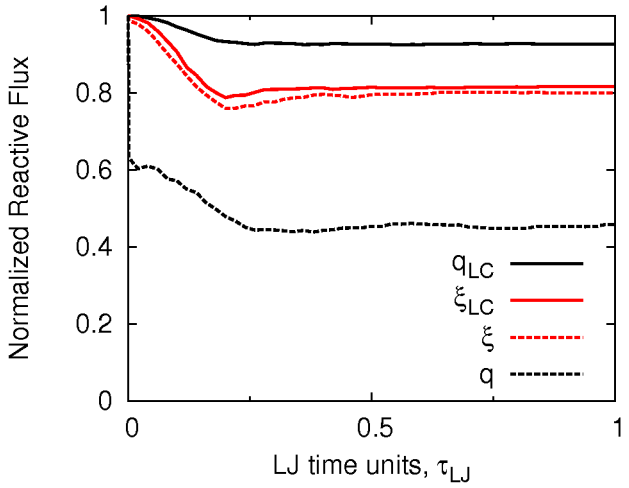

Figure 3.3. The normalized reactive flux correlation function for the 0-isosurface of each coordinate.

Additionally, we computed κ for each trial coordinate from the normalized reactive flux correlation function (see Table 3.1 and Fig. 3.3). In each case, recrossing occurs during the first 0.5 τ_{LJ} and thereafter the reactive flux plateaus. The highest κ is $\kappa[q_{LC}] = 0.93$, confirming the iLMax result that q_{LC} is the best coordinate. The lowest transmission coefficient is $\kappa[q] = 0.45$; q also scored the lowest $\Delta \ln L$. Most of the drop in $\kappa[q]$ occurs in the first 0.001 τ_{LJ}, a consequence of the uncertainty in locating D and A, and therefore the $q = 0$ dividing surface, by the energy minimization technique. After 0.001 τ_{LJ}, the drop in $\kappa[q]$ is comparable to the drop in κ seen for other trial coordinates. The uncertainty in locating $q = 0$ matters most when the hopping particle H is very close to the dividing surface. In contrast, $\kappa[\xi] = 0.80$ and $\kappa[\xi_{LC}] = 0.81$ are similar, highlighting that the dependence of ξ on the precise locations of D and A is relatively minor. The difference between $\kappa[q_{LC}]$ and $\kappa[q]$ is larger than the difference between $\kappa[q_{LC}]$ and $\kappa[\xi_{LC}]$, showing that the accuracy with which D and A are located has a greater impact than the form of the trial coordinate.

We anticipate that the degree to which $\kappa[q_{LC}]$ improves over $\kappa[\xi]$ (or, equivalently, $\kappa[\xi_{LC}]$) will be system dependent. In particular, we note two studies that computed $\kappa[\xi]$ for

vacancies in different crystal lattices. First, in an FCC crystal the hopping particle must pass through a gate composed of 4 nearest-neighbor particles. The largest void space between gate particles lies on the straight line path from D to A, as is also the case for the 2D trigonal crystal. In an FCC crystal of LJ particles, $\kappa[\xi] = 0.91$ at $k_B T/\epsilon = 0.505$ [21]. Although we expect $\kappa[q] > \kappa[\xi]$, $\kappa[\xi]$ is already close to 1. Any improvement that q provides will not make a dramatic difference. Second, in a rock-salt lattice, an anion that hops into an adjacent vacancy can either pass directly between two occupied cation sites or through the larger off-center void space. ξ does not characterize off-center motion. Accordingly, for oxygen hopping in cobalt oxide crystals $\kappa[\xi] \approx 0.3$ over a wide range of temperatures, 995 – 1760 K [22]. Since $\kappa[\xi]$ is less than 1, the $\xi = 0$ surface is recrossed many times and we anticipate that $\kappa[q_{LC}]$ will be a significant improvement over $\kappa[\xi]$.

3.5 Conclusions

Permutation shooting was applied to vacancy migration in a 2D trigonal crystal of Lennard-Jones particles. Using inertial likelihood maximization to analyze the TPS data, we identified the reaction coordinate and easily computed the rate constant. Moreover, likelihood maximization was able to discern good coordinates from poor ones even when our intuition did not anticipate a big difference. A numerical problem with optimization procedures for describing the donor to acceptor direction was revealed that helped improve estimates of hopping rates and transmission coefficients. Of the two proposed order parameters—position along the straight line path between donor and acceptor sites ξ and the bipolar distance q—the better coordinate is q.

3.6 References

1. Fahey, P. M.; Griffin, P.; Plummer, J. "Point defects and dopant diffusion in silicon." *Rev. Mod. Phys.* **1989**, *61*, 289.
2. Michel, A.; Rausch, W.; Ronsheim, P.; Kastl, R. "Rapid annealing and the anomalous diffusion of ion implanted boron into silicon." *Appl. Phys. Lett.* **1987**, *50*, 416-418.
3. Munro, L. J.; Wales, D. J. "Defect migration in crystalline silicon." *Phys. Rev. B* **1999**, *59*, 3969.
4. Chatti, I.; Delahaye, A.; Fournaison, L.; Petitet, J.-P. "Benefits and drawbacks of clathrate hydrates: a review of their areas of interest." *Energy Conversion and Management* **2005**, *46*, 1333-1343.
5. Lee, H.; Lee, J.-w.; Do Youn Kim, J. P.; Seo, Y.-T.; Zeng, H.; Moudrakovski, I. L.; Ratcliffe, C. I.; Ripmeester, J. A. "Tuning clathrate hydrates for hydrogen storage." *Nature* **2005**, *434*, 743-746.
6. Peters, B.; Zimmermann, N. E.; Beckham, G. T.; Tester, J. W.; Trout, B. L. "Path sampling calculation of methane diffusivity in natural gas hydrates from a water-vacancy assisted mechanism." *J. Am. Chem. Soc.* **2008**, *130*, 17342-17350.
7. Liu, X.; Mayer, M. T.; Wang, D. "Understanding ionic vacancy diffusion growth of cuprous sulfide nanowires." *Angew. Chem. Int. Ed.* **2010**, *49*, 3165-3168.
8. Phillips, R.: *Crystals, defects and microstructures: modeling across scales*; Cambridge University Press, 2001.
9. Mars, P.; Van Krevelen, D. W. "Oxidations carried out by means of vanadium oxide catalysts." *Chem. Eng. Sci.* **1954**, *3*, 41-59.
10. Zinkle, S.; Ghoniem, N. "Operating temperature windows for fusion reactor structural materials." *Fusion Eng. Des.* **2000**, *51*, 55-71.
11. Odette, G.; Alinger, M.; Wirth, B. "Recent developments in irradiation-resistant steels." *Annu. Rev. Mater. Res.* **2008**, *38*, 471-503.
12. Plimpton, S. "Fast parallel algorithms for short-range molecular dynamics." *J. Comput. Phys.* **1995**, *117*, 1-19.
13. Geslin, P.-A.; Ciccotti, G.; Meloni, S. "An observable for vacancy characterization and diffusion in crystals." *J. Chem. Phys.* **2013**, *138*, 144103.
14. Weeks, J. D.; Chandler, D.; Andersen, H. C. "Role of repulsive forces in determining the equilibrium structure of simple liquids." *J. Chem. Phys.* **1971**, *54*, 5237-5247.
15. Bennett, C. H. "Exact defect calculations in model substances." *Diffusion in Solids: Recent Developments* **1975**, 73-113.
16. Peters, B. "Inertial likelihood maximization for reaction coordinates with high transmission coefficients." *Chem. Phys. Lett.* **2012**, *554*, 248-253.
17. Radhakrishnan, R.; Schlick, T. "Biomolecular free energy profiles by a shooting/umbrella sampling protocol, "BOLAS"." *J. Chem. Phys.* **2004**, *121*, 2436-2444.
18. Nose, S. "A unified formulation of the constant temperature molecular dynamics methods." *J. Chem. Phys.* **1984**, *81*, 511-519.
19. Hoover, W. G. "Canonical dynamics: equilibrium phase-space distributions." *Phys. Rev. A* **1985**, *31*, 1695.
20. Chandler, D. "Statistical mechanics of isomerization dynamics in liquids and the transition state approximation." *J. Chem. Phys.* **1978**, *68*, 2959-2970.

21. Paci, E.; Ciccotti, G. "Vacancy migration rates by molecular dynamics with constraints." *J. Phys.: Condens. Matter* **1992**, *4*, 2173.
22. Gillan, M.; Harding, J.; Tarento, R.-J. "Calculation of defect migration rates by molecular dynamics simulation." *J. Phys. C* **1987**, *20*, 2331.

Chapter 4

Transmission Coefficients, Committors, and Solvent Coordinates in Ion-pair Dissociation

Solvent effects are well known to influence the rates of reactions [1,2]. A solvent can alter the reaction rate by changing the activation energy and the activation entropy [3,4]. A solvent can also alter the dynamics of barrier crossing and thereby change the kinetic prefactor of a reaction rate constant [5-8]. Unfortunately, experimental attempts to change the dynamics also tend to change the activation barriers. For example, increasing the molecular weight of an alkane solvent to increase solvent viscosity makes small differences in the barrier which are exponentially magnified in the rate constant [9].

Even theoretical treatments can conflate equilibrium and dynamic effects. The true rate constant k can be obtained from a dynamic correction κ to the transition state theory (TST) rate constant k_{TST}, which is computed entirely from equilibrium properties [10],

$$k = \kappa[q]\, k_{TST}[q] \tag{4.1}$$

k is independent of the choice of reaction coordinate q [11]. So that their product remains constant, κ and k_{TST} must change inversely with a change in dividing surface $q(\mathbf{x}) = q_{\ddagger}$ [12-14]. In the dividing surface definition q is computed from atomic coordinates **x** and $q_{\ddagger}$ is the

value of the reaction coordinate at the dividing surface. $\kappa[q]$ is the plateau value of the reactive flux correlation function [15]

$$\kappa[q] = \frac{\langle \dot{q}(0)\, \theta\big(q(t_{plateau}) - q_{\ddagger}\big)\, \delta\big(q_{\ddagger} - q(0)\big)\rangle}{\langle \dot{q}\, \theta(\dot{q})\, \delta\big(q_{\ddagger} - q\big)\rangle} \tag{4.2}$$

where $\theta(q(t)) = 1$ if q is on the product (B) side of the barrier at time t and $\theta = 0$ otherwise, and t_{mol} is the molecular relaxation time. In Eq. (4.2), the notation $q(\mathbf{x}(t))$ has been simplified to $q(t)$. Apart from tunneling [16-18], no dynamical correction is needed, i.e. $\kappa = 1$, when reactant-initiated trajectories that cross the dividing surface $q(\mathbf{x}) = q_{\ddagger}$ continue on to products without recrossing. If some trajectories do recross the dividing surface, $k_{\mathrm{TST}}[q]$ will overestimate k [12] with dynamic correction $\kappa[q] < 1$. However, recrossing is not necessarily dynamic in origin. One can show, e.g. for Langevin dynamics on a simple two-dimensional potential energy surface (PES), that recrossing can result from intrinsic friction in the dynamics [5,6], from a poorly chosen dividing surface [19,20], or from a combination of these factors. In an atomic system, friction can emerge from solvent or other degrees of freedom projected away en route to the potential of mean force [2,21] or from electronic friction as experienced by adsorbates bound to a metal [22]. There are many ways to choose the wrong reaction coordinate before projection, with each wrong way resulting in a poorly chosen dividing surface. In a high dimensional system, it is difficult to conclusively determine whether recrossing emerges from some intrinsic friction, or from a poorly chosen reaction coordinate and dividing surface. Since the choice of q changes the apparent dynamic and equilibrium contributions to k, k_{TST} should ideally be minimized to ensure that κ is indicative of true dynamical friction.

In Variational Transition State Theory (VTST) [14,23-25], the dividing surface $q(\mathbf{x}) = q_{\ddagger}$ is optimized to minimize recrossing (and thereby k_{TST}) by changing the value $q_{\ddagger}$ or the

reaction coordinate $q(\mathbf{x})$, i.e., by changing the coordinate system. VTST is particularly successful when the bottleneck is a single, high saddle on the PES. Within a harmonic approximation to the saddle region, a perfect dividing surface with no recrossing can be obtained from normal mode analysis [26-30]. This procedure leads to the well-known and widely used harmonic TST [26,27]. For gas phase chemistry and reactions on solid surfaces where the bottleneck is a single saddle on the PES, the transmission coefficients for the dividing surface of harmonic TST is usually very close to unity [23,31,32]. Note that in contrast to many rare events methods for more complicated reactions [33,34], the highly successful harmonic TST optimizes the reaction coordinate before computing free energy barriers and prefactors. VTST has also been used for more complicated reactions, e.g. in enzymes and in solution, but these applications require *a priori* specification of a small subset of important coordinates [23,35-37]. For reactions in which the solvent is involved in truly complex and unknown ways, a complete space VTST calculation has yet to be demonstrated.

As an alternative to VTST, likelihood maximization [19,30,38] makes use of Transition Path Sampling (TPS) [39] data that is collected without *a priori* knowledge of pre-selected coordinates. The original version of likelihood maximization (oLMax) [19,38] identifies coordinates that accurately predict the committor probability [40-45] $p_B(\mathbf{x})$, i.e. the fraction of trajectories initiated from configuration $\mathbf{x}$ with Boltzmann distributed momenta that terminate as products (B) rather than reactants (A). However, the newer inertial likelihood maximization (iLMax) [30] also uses the velocity along the trial coordinates to ensure that coordinates accurately predict the committor *and* maximize the transmission coefficient κ.

Neither VTST nor iLMax can find a dividing surface with no recrossings unless such a surface exists. In the past it was difficult to prove or disprove the existence of exact dividing surfaces. This difficulty always left the possibility that a complete converged VTST calculation or an exhaustive screening of trial coordinates with iLMax might give the exact rate with no further dynamical corrections. We have recently developed an existence test for no-recrossing surfaces based on rejected shooting moves in TPS [46]. This paper introduces an additional existence test, this one utilizing the committor, that also provides insight into the nature of solvent-induced recrossing. We apply the new test to NaCl dissociation in water. Moreover, we employ both oLMax and iLMax and contrast the coordinates identified by each by computing the committor distribution [47] and κ from Eq. (4.2). We gain insight into the reaction mechanism by investigating transition state configurations and computing free energy landscapes.

4.1 Methods

Simulations of one NaCl ion pair and 394 water molecules were conducted under isothermal isobaric (NPT) conditions at 300 K and 1 bar with periodic boundary conditions. The simulation cell was a cubic box 22.9 Å long, on average. Water was modeled using the TIP3P parameters [48]. The interaction energy between particles includes both Lennard-Jones (LJ) and Coulomb contributions. LJ parameters from the OPLS force field were used for the ions [49] .Parameters for cross-interactions were determined using geometric combination rules. Electrostatic interactions were calculated using Particle Mesh Ewald [50] with 28 grid points in each direction. Dynamics were simulated in NAMD [51] using a 0.25 fs time step. Langevin dynamics were performed with a $1/\tau = 0.1$ ps^{-1} damping

coefficient. The Nosé-Hoover barostat [52] was used for pressure control with a period of 100 fs and a decay time scale of 50 fs.

Reactant and product basin definitions for committor analysis and TPS were established by projecting the free energy onto the ion-pair distance r_{ion}. Free energies were computed by umbrella sampling using a harmonic biasing potential in r_{ion}. Thirty-seven windows were used, each with a different central ion-pair distance parameter, r_w, ranging from 2.5 to 9.75 Å. Each trajectory was 3 ns long and configurations were saved every 20 fs, resulting in a total of 5.5 million saved configurations. The free energy $F(r_{ion})$ was then constructed using the Weighted Histogram Analysis Method (WHAM) [53]. Convergence was verified by comparison to the work of Fennell et al. [49] Free energy surfaces projected onto other order parameters were computed by reweighting.

The free energy of the committor, $F(p_B)$, was computed using a subset of the sampled configurations. We computed p_B-estimates using 20 trajectories per estimate for 10,500 configurations on the barrier, drawn from seven umbrella sampling trajectories (3.1 Å $\leq r_w \leq$ 4.25 Å). The average p_B for the $r_w = 3.1$ and 4.25 Å trajectories was 0.03 and 0.96, respectively. Therefore, the free energy of the stable basins was estimated by setting p_B to 0 or 1 for configurations from umbrella trajectories below or above this range, respectively. Although 10,500 configurations is fewer than the 5.5 million configurations used for other order parameters, the barrier $\Delta F(r_{ion})$ is only 0.1 $k_B T$ different when estimated from this reduced subset of configurations.

Basins should include typical thermal fluctuations within stable states, but should not encroach on the transition path region. These conflicting requirements necessitated development of variable path length Aimless Shooting which allows reactant and product

basins to be defined far apart from each other so as not to interfere with mechanistic reaction coordinate analysis.

We defined basin boundaries at the local free energy minima on either side of the barrier. The contact state is defined as $r_{ion} < 2.7$ Å and the dissociated state is $r_{ion} > 5.3$ Å. Because the interionic forces are repulsive at short distances, high-energy trajectories that reach $r_{ion} = 2.7$ Å could be redirected back across the barrier. Therefore, we also require the kinetic energy along r_{ion}, averaged over the next 300 fs (4 equilibrium oscillatory periods), be less than $0.25\ k_B T$. We performed $N_R = 24{,}000$ Aimless Shooting moves with two candidate shooting points separated by $\delta t = 40$ fs. Trajectories were on average 2.9 ps long, with a standard deviation of 1.9 ps.

Solvent coordinates were optimized using both oLMax and iLMax. For both methods, we used a reaction coordinate model $q(\mathbf{z}) = q_T(\mathbf{z}) + c_A\, \ddot{q}_T(\mathbf{z})$ where the trial coordinate $q_T(\mathbf{z}) = c_0 + \mathbf{c}\cdot\mathbf{z}$ is a linear combination of $\mathbf{z} = (z_1, z_2, \ldots, z_M)$ order parameters with coefficients $\mathbf{c} = (c_1, c_2, \ldots, c_M)$. The order parameters $\mathbf{z}$ in $q_T(\mathbf{z})$ are for the shooting point configuration (t=0) of each trajectory. Accelerations $\ddot{\mathbf{z}}$ for the shooting points were calculated by central finite difference using configurations ±1 NVE timestep. The NVE timestep is needed to eliminate any stochastic effects of the thermostat or barostat from $\ddot{\mathbf{z}}$. We used the error function to model the committor probability $p_B = (1+\mathrm{erf}[q(\mathbf{z})]) / 2$ for oLMax and the reaction probability $p_{RX} = (1+\mathrm{erf}[q(\mathbf{z}) + c_V\, \dot{q}(\mathbf{z})]) / 2$ for iLMax. When q includes the accelerations $\ddot{\mathbf{z}}$, the third time derivatives of $\mathbf{z}$, called the jerks, are also required as generalized velocities of the acceleration (force) component. The velocities $\dot{\mathbf{z}}$ and jerks needed for $\dot{q}(\mathbf{z})$ were computed by finite differencing $\mathbf{z}$ and $\ddot{\mathbf{z}}$, respectively, for the t=0 and t=0.25 fs configurations. The coefficients c_0, $\mathbf{c}$, c_A and c_V are determined when maximizing

the likelihood. Linear combinations of up to $M = 3$ order parameters were tested. The Bayesian information criteria $\delta L_{\min} = \frac{1}{2} \ln(N_R)$, which is 5.0 in this work, helps to discriminate between significant and insignificant coordinate improvements.

Initial configurations for the committor analysis [40,41] and the reactive flux correlation function [15] are from the equilibrium distribution in a narrow interval at the putative transition state. We collected configurations $r_{ion} = 3.7 \pm 0.02$ Å from the $r_w = 3.7$ Å umbrella sampling trajectory. All saved configurations were at least 2 ps apart. For coordinates identified by likelihood maximization, configurations $q(\mathbf{z}) = 0 \pm 0.02$ were also selected from saved umbrella sampling configurations, sampled according to weights from the $q(\mathbf{z}) = 0$ isosurface of the free energy $F[r_{ion}, q(\mathbf{z})]$. All starting ensembles consisted of 1000 configurations, except for p_B which contained 100.

For committor distributions, p_B was estimated using 20 trajectories per configuration. After deconvolution, the mean and standard deviation of the intrinsic distribution $p(p_B \mid q)$ provide a beta distribution model which is shown along with the raw histogram [54]. The reactive flux correlation function was computed from forward and time-reversed trajectories with reaction coordinates evaluated every 20 fs. The velocity along each coordinate $dq/dt|_{q=q^*}$ was approximated by finite difference using $\Delta t = 0.25$ fs. For $\kappa[p_B]$, Δp_B was estimated using a longer time interval $\Delta t = 20$ fs. The long time interval for estimating dp_B/dt is required because of statistical uncertainty in the value of p_B, as shown in Fig. 4.1. p_B-estimates were computed using up to $N = 1000$ trajectories, for which the uncertainty δp_B is ± 0.016.

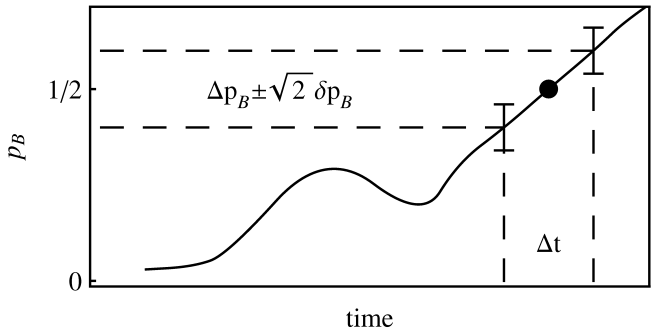

Figure 4.1. Schematic of a forward and time-reversed trajectory pair used to compute $\kappa[p_B]$. The derivative $dp_B/dt|_{pB = 1/2}$ is estimated by central finite difference. We found that $\Delta t = 20$ fs gave small errors (see Appendix B).

4.2 Testing the No-Recrossing Hypothesis

The existence of a no-recrossing surface can be potentially falsified using the committor p_B. To illustrate the connection between p_B and recrossing, consider forward and time-reversed (backward) trajectories initiated from any dividing surface, not just one free of recrossing. These forward and backward trajectory pairs can be grouped into three types:

- **Non-recrossing trajectories.** For each momentum the forward and backward trajectories go directly to opposite states from the dividing surface with no recrossing. If all trajectories initiated from a dividing surface are non-recrossing, κ will equal 1 and p_B will be ½ for all states on the dividing surface.
- **Recross in one direction.** Trajectories for which both directions reach A (B) contribute 0 (1) to a p_B-estimate. A dividing surface with any trajectories of this type must have $\kappa < 1$.
- **Recross in both directions.** Trajectories that recross the dividing surface in both directions contribute ½ to p_B, but cause $\kappa < 1$. Trajectories of this type demonstrate that optimizing for p_B does not necessarily optimize κ.

As noted in the introduction, a dividing surface with no recrossing and $\kappa = 1$ must be composed entirely of $p_B = ½$ states.

A no-recrossing surface, if it exists, also divides configuration space into a basin of attraction for A ($p_B \leq ½$) and a basin of attraction for B ($p_B \geq ½$), as shown in Fig. 4.2. If the no-recrossing surface exists, then a trajectory that begins in A and crosses into the $p_B > ½$ region must end in B. By initiating trajectories from $p_B = ½$ configurations—exactly the procedure for calculating $\kappa[p_B]$—we can screen for non-reactive trajectories that prove the non-existence of a recrossing-free dividing surface.

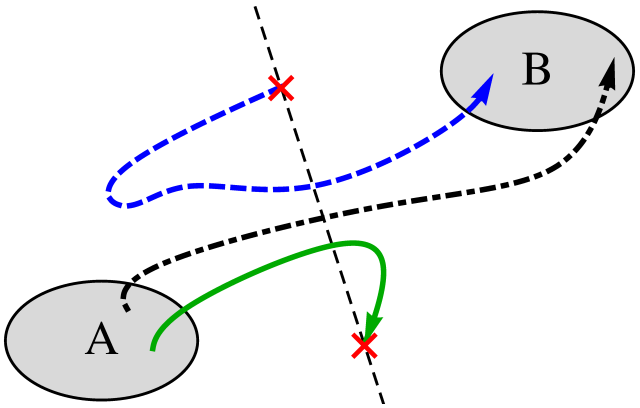

Figure 4.2. Schematic of system with two stable states A and B for which a recrossing-free dividing surface (black, dashed) exists. By definition, only A→B transition paths (black, dot-dashed) cross the dividing surface and therefore $p_B = ½$ for configurations on the dividing surface. A trajectory that begins in A and crosses into the $p_B > ½$ region cannot also end in A (green). Similarly, a B→B trajectory (blue, dashed) cannot cross into $p_B < ½$.

4.3 Ion-pair dissociation

4.3.1 History and value as a test system

Ion-pair dissociation is a frequent test case for investigating solvent dynamics through theory and simulation [41,55-60]. When two ionic solutes encounter and form a contact pair in solution, the solvation shells must change to accommodate the new solute configuration. Solvent rearrangements are thought to be critical for ionic crystal growth rates [61,62], the formation of contacts between charged residues in proteins [63,64] and for electron transfer rates which depend on both the distance between donor and acceptor ions [65,66] and the solvent environment [67]. From a purely theoretical perspective, ion-pair dissociation is

somewhat different from the S_N2 reactions where VTST has been successful. Specifically, S_N2 reactions have a barrier with or without the solvent present. The solvent can then be viewed as a small perturbation on the solute reaction coordinate. In contrast, ion-pair association is a barrierless process in the absence of solvent. Therefore, the solvent completely changes the landscape on which ion-pair dissociation occurs [57,68].

Contact ion pairs can be distinguished from solvent-separated ion pairs according to the distance between the ions, $r_{ion} = |\, \boldsymbol{r}_{cation} - \boldsymbol{r}_{anion} \,|$. Fig. 4.3 shows the free energy $F(r_{ion})$ for a NaCl ion pair in TIP3P water. The barrier to dissociation peaks at $r_{ion} = 3.7$ Å and 5.2 k_BT, suggesting that contact ion pairs are only narrowly metastable.

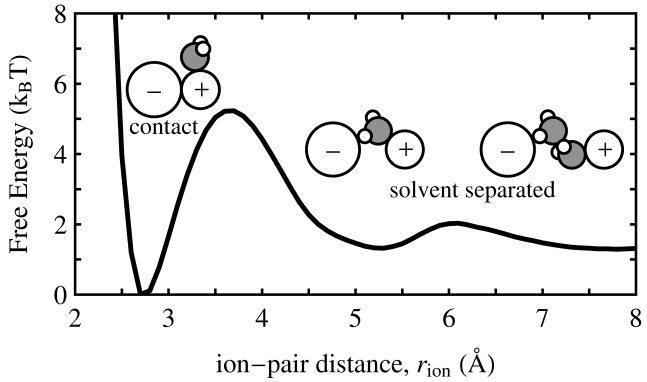

Figure 4.3. Free energy $\beta F(r_{ion})$ of the aqueous NaCl system as a function of ion-pair distance r_{ion}. Along this coordinate, a barrier of 5.2 k_BT separates the contact ion pair from the local maximum of $F(r_{ion})$ at $r_{ion} = 3.7$ Å. For a single ion pair without periodic boundary conditions, the free energy diverges as -2 $\ln(r_{ion})$ in the limit $r_{ion} \to \infty$.

From the reactive flux correlation function, $\kappa[r_{ion}] = 0.29 \pm 0.02$ (see Fig. 4.5, thin line), in general agreement with previously published results for other ion and water force fields [41,57,58]. Geissler et al. [41] computed the distribution of committors $p(p_B \mid r_{ion})$ for configurations constrained to free energy peak in $F(r_{ion})$. For a good coordinate q, the $q_‡$ dividing surface should identify transition states ($p_B = ½$), so $p(p_B \mid q)$ should be peaked at $p_B = ½$ with a small standard deviation σ. However, $p(p_B \mid r_{ion})$ is bimodal (see Fig. 4.7(a))

with most configurations at the barrier top committed to the contact basin ($p_B = 0$) or solvent-separated basin ($p_B = 1$) and few transition states. Their analysis, confirmed by subsequent investigations [69,70], shows that some as-yet unidentified solvent variables are important in the reaction coordinate. Truhlar and Garrett [59] demonstrated that a bimodal $p(p_B \mid r_{ion})$ can emerge when r_{ion} is coupled harmonically with even a single solvent coordinate, and further suggested that recrossing would be completely eliminated by finding that coordinate.

4.3.2 Dynamics of the Committor

The thermal ensemble of $p_B = ½$ configurations needed to calculate $\kappa[p_B]$ was sampled by first computing the free energy $F(p_B)$. Fig. 4.4 shows the $\Delta F(p_B)$ barrier increases to 7.2 k_BT relative to $\Delta F(r_{ion}) = 5.2\ k_BT$. Interestingly, the top of $F(p_B)$ is almost, but not completely, flat as anticipated from the backward Kolmogorov equation [42].

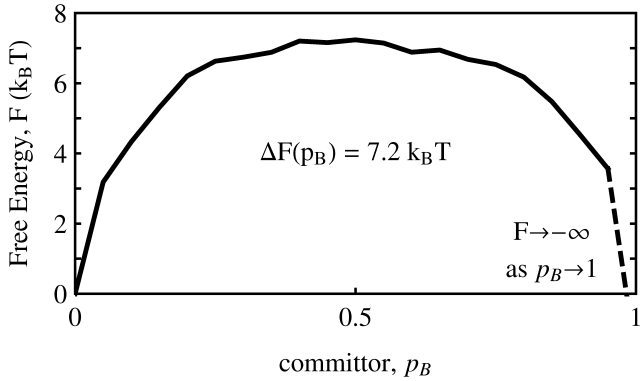

Figure 4.4. Free energy $\beta F(p_B)$ showing the dissociation barrier is 7.2 k_BT. In principle, the free energy diverges as $p_B \rightarrow 1$ since all configurations with $r_{ion} > 5.3$ Å belong to state B.

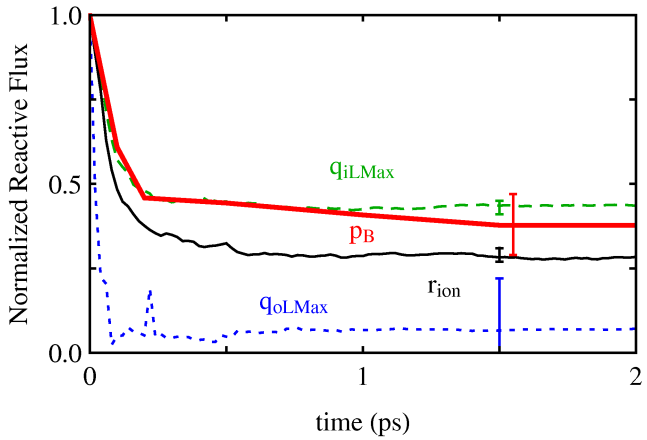

Figure 4.5. The normalized reactive flux correlation function for the dividing surfaces $p_B = 1/2$ (thick red), $r_{ion} = 3.7$ Å (black), $q_{iLMax} = 0$ (green, dashed) and $q_{oLMax} = 0$ (blue, dotted). Error bars are shown for the plateau value κ. The larger uncertainty for $\kappa[p_B]$ results from using 100 trajectories to calculate the reactive flux, as opposed to 1000 trajectories for other coordinates.

To calculate $\kappa[p_B]$, we initiated 100 trajectories with Boltzmann distributed momenta, each from a different configuration with $p_B = 0.5 \pm 0.016$. The transmission coefficient is $\kappa[p_B] = 0.38 \pm 0.09$ (see Fig. 4.5, thick line) compared to $\kappa[r_{ion}] = 0.29 \pm 0.02$. The uncertainty is larger in $\kappa[p_B]$ than in $\kappa[r_{ion}]$ due to the number of trajectories used (Appendix B). However, it is clear that $\kappa[p_B] < 1$. Therefore, we can conclude that friction-induced recrossing in NaCl dissociation persists even for the $p_B = ½$ dividing surface.

Additionally, we screened these 100 trajectories for non-reactive cases to determine whether $\kappa = 1$ is possible for this system. Because of the uncertainty in calculating p_B, we only assign a degree of statistical confidence that a trajectory has crossed $p_B = ½$. For example, we are 99.7% confident that an A→A trajectory which contains a configuration with $p_B > 0.548$ crossed into state B. Twenty-three of the 100 trajectories met or exceeded this standard, two of which are shown in Fig. 4.6. These provide compelling evidence that a recrossing-free surface for ion-pair dissociation does not exist. The remainder of the manuscript focuses on learning *how* the solvent dynamics create inextricable recrossing.

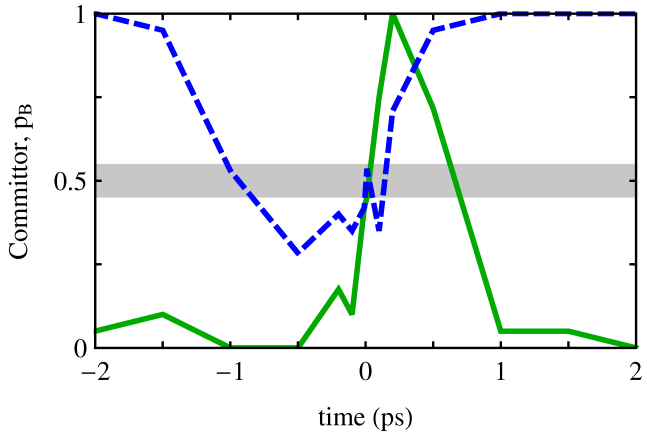

Figure 4.6. The committor p_B as a function of time for two non-reactive trajectories initiated from the $p_B = ½$ isosurface. The A→A trajectory (green) and B→B trajectory (blue, dashed) prove the non-existence of a recrossing-free surface by crossing the $p_B = ½$ isosurface. The 99.7% confidence interval $\pm 3\delta p_B = \pm 0.038$ is shaded around $p_B = 0.5$.

4.3.3 Solvent Coordinates

The increase in $\kappa[p_B]$ from $\kappa[r_{ion}]$ suggests that recrossing can be at least partially eliminated by optimization of the reaction coordinate. However, p_B is essentially a statistical descriptor of the dynamics that conveys no mechanistic information about the common *physical* characteristics of the transition states [42,63]. By contrast, reaction coordinates based on concrete physical characteristics of the atomic configuration are inexpensive to compute and convey valuable insight about the reaction mechanism. Besides r_{ion}, all other order parameters must involve the solvent. Formally, a solvent coordinate should depend only on the instantaneous density of solvent molecules or atoms. For both oLMax and iLMax, we use a trial reaction coordinate model $q_T(\mathbf{z})$ that is a linear combination of $\mathbf{z} = (z_1, z_2, \ldots, z_M)$ order parameters with M = 1, 2, or 3. We include 50 candidate variables that depend only on the local density of solvent molecules or atoms relative to the ion pair, such as 1st and 2nd shell coordination numbers, water density between the ions, energy gaps

and other electrostatic coordinates. A full list of candidate variables considered, many taken from previous studies [41,56,58,71-74], is included in Appendix C.

Additionally, we examined the regimes of Grote-Hynes (GH) theory [6] to extend our list of trial solvent coordinates. GH theory successfully predicts $\kappa[r_{ion}]$ by invoking a non-Markovian time-dependent friction $\eta(t)$ to model the dynamics of r_{ion} [56,57]. In GH theory, κ depends on the frequency-dependent Laplace transform of the friction, $\tilde{\eta}(\lambda)$, on the range $0 < \lambda < \omega_b$, where ω_b is the frequency of the solvent-equilibrated free energy barrier. Markovian friction is constant over this range, $\tilde{\eta}(\omega_b) \approx \tilde{\eta}(0)$. Non-Markov effects are important for $\tilde{\eta}(\omega_b) \ll \tilde{\eta}(0)$. In evaluating $\tilde{\eta}(\omega_b)$, the timescale ω_b^{-1} is intrinsically compared to the friction correlation time τ_c. τ_c is the ratio of the zero-frequency friction to the initial time-dependent friction, $\tilde{\eta}(0) / \eta(0)$ [75], and is one way to characterize the spectrum of solvent response times embedded in $\eta(t)$. The non-Markov regime $\tilde{\eta}(\omega_b) \ll \tilde{\eta}(0)$ applies when the solvent response times are slow, $\tau_c \gg \omega_b^{-1}$. Our results show, in agreement with previous results [56,57], that friction in ion-pair dissociation is non-Markovian. We find $\tilde{\eta}(\omega_b) / \tilde{\eta}(0) = 0.14$ and $\tau_c = 160$ fs is larger than $\omega_b^{-1} = 60$ fs.

There are two relevant non-Markovian regimes of GH theory, differentiated by the strength of the initial time-dependent friction $\eta(0)$.

- **Nonadiabatic regime, $\eta(0) \ll \omega_b^2$.** When the bath is slow but weakly coupled to the solute coordinate, the bath modes, unable to evolve during the barrier crossing, are frozen and therefore exert a nearly constant force on the solute coordinate.
- **Dynamic caging regime, $\eta(0) \gg \omega_b^2$.** When the bath is slow and strongly coupled to the solute coordinate, only a narrow range of solute coordinate values are allowed at

the barrier top. The solute coordinate vibrates in this range, or cage, until the bath modes slowly evolve toward the reactant or product basin.

For ion-pair dissociation, $\eta(0) / \omega_b^2 = 5$. With less than an order of magnitude separation, elements of either the nonadiabatic or dynamic caging regimes may be present. We developed trial coordinates pertaining to each.

In the nonadiabatic regime, trajectories recross the dividing surface because the force on the solute coordinate from the bath modifies the shape of the barrier. For ion-pair dissociation, the direct Coulombic force between the ions is modified by the additional force exerted by the solvent. An analytic derivation for the solvent force is given in Ref. [57]. We included the total force on the ion-pair distance, f_{rion}, as a trial coordinate and generalized this idea by including an acceleration term $c_A \ddot{q}_T(\mathbf{z})$ in the committor model for all trial reaction coordinates. The parameter c_A is adjusted when maximizing the likelihood, so adding the acceleration of $q_T(\mathbf{z})$ has the same effect as testing coordinates in combination with the forces that act on them.

In the dynamic caging regime, a trajectory projected onto the solute coordinate will oscillate [76]. The minimum energy solute coordinate allowed by the current bath modes will be a better indicator of reaction progress than the instantaneous reaction coordinate. Therefore, for each shooting point, we calculate a reference configuration by holding the water coordinates fixed and moving the ion positions to minimize the potential energy. Order parameters describing this reference configuration are labeled with a subscript 'opt', e.g. the optimized ion-pair distance r_{opt} corresponds to the minimum energy ion-pair distance within the frozen solvent configuration. Other examples are described in Appendix C. Anticipating that the motion of hydrogen atoms may be faster than the solute motion, we tested similar

order parameters by also optimizing, in addition to the ion positions, the orientations of water molecules coordinated to either ion.

4.3.4 Original Likelihood Maximization (oLMax)

Likelihood results are shown in Table 4.1 for the solute order parameter r_{ion} and the best combinations with other solvent coordinates. Log-likelihoods for each coordinate are reported relative to $\ln L[r_{\mathrm{ion}}]$ in δL_{min} increments,

$$\Delta \ln L\,[q] = \frac{\ln L[q] - \ln L[r_{\mathrm{ion}}]}{\delta L_{\mathrm{min}}} \tag{4.3}$$

where the absolute $\ln L[r_{\mathrm{ion}}]$ = -10,239 and where δL_{min} for this work is 5.043 (see Methods). In each case, the predicted $p_{\mathrm{B}} = ½$ dividing surface from likelihood maximization is $q = 0$, which for r_{ion} gives $r_{‡} = -c_0 / c_1 = 3.84$ Å, a distance slightly larger than the location of the free energy barrier top 3.7 Å. Table 4.1 shows that linear combinations of r_{ion} with either f_{rion} or r_{opt} improves the reaction coordinate by several δL_{min}, indicating that elements of both the nonadiabatic and dynamic caging regimes are relevant to the dynamics. Furthermore, the linear combination $q_{\mathrm{T}}(r_{\mathrm{ion}}, f_{\mathrm{rion}})$ scores the same $\ln L$ as $q_{\mathrm{T}}(r_{\mathrm{ion}})+c_{\mathrm{A}}\ \ddot{q}_{\mathrm{T}}(r_{\mathrm{ion}})$, validating our methodology to use finite-differenced accelerations in lieu of forces. The best among our trial coordinates from oLMax, with $\Delta\ln L = 185$, is a function of r_{ion}, r_{opt}, and the number of waters jointly coordinated to both ions N_{B}. We label this coordinate q_{oLMax}.

q_{oLMax} is an improvement over r_{ion} as measured by the committor histogram test, but not according to the reactive flux correlation function. Fig. 4.7(b) shows the committor distribution $p(p_{\mathrm{B}} \mid q_{\mathrm{oLMax}})$ with a single maximum near $p_{\mathrm{B}} = ½$, in contrast to the bimodal $p(p_{\mathrm{B}} \mid r_{\mathrm{ion}})$ in Fig. 4.7 (a). However, rather than increasing from $\kappa[r_{\mathrm{ion}}] = 0.29$, Fig. 4.5 (dotted curve) shows $\kappa[q_{\mathrm{oLMax}}]$ decreases to 0.07. Therefore, by including solvent modes in the

reaction coordinate, κ became *lower* even while the reaction coordinate became a better predictor of p_B.

Table 4.1. Select Results from Original Likelihood Maximization[a]

$q(\mathbf{z})$[b]	$\Delta \ln L$[c]
$1.5\, r_{ion} - 5.7$	0
$1.5\, r_{ion} + 0.034\, f_{rion} - 5.6$	+35
$1.5\, r_{ion} + 112\, \ddot{r}_{ion} - 5.6$	+35
$0.62\, r_{ion} + 0.87\, r_{opt} - 5.8$	+43
$0.64\, r_{ion} + 0.50\, N_B + 18\, r_{opt}^{-2} - 1.8$	+185

[a]The best coordinate is indicated by the highest likelihood.
[b]Units are: r_{ion} (Å), $\ddot{r}_{ion}$ (Å/fs^2), r_{opt} (Å), f_{rion} (kcal/mol/Å), N_B (unitless).
[c]Likelihoods $\Delta \ln L$ are reported in δL_{min} increments relative to that for r_{ion}.

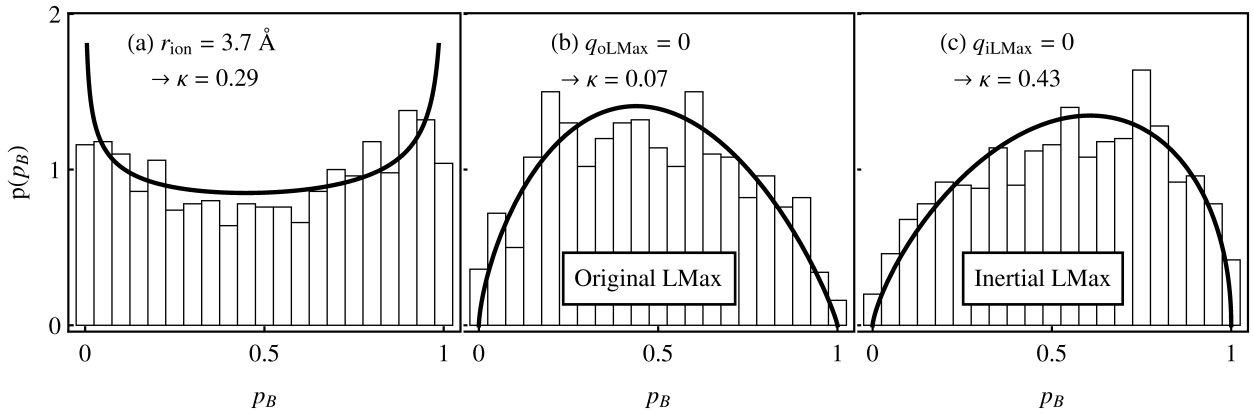

Figure 4.7. Beta-distribution models of $p(p_B)$ (line) fit to the histogram of p_B-estimates (bars) for the constrained ensembles **(a)** $r_{ion} = 3.7$ Å, **(b)** the original likelihood maximization coordinate $q_{oLMax} = 0$, and **(c)** the inertial likelihood maximization coordinate $q_{iLMax} = 0$. Reaction coordinate accuracy increases as the standard deviation σ of the distribution decreases.

This conflicting result was foreshadowed in section 4.2, where we examined the forward-backward trajectory pairs initiated from the dividing surface. Analyzing individual trajectories confirms that the fraction of the reactive flux associated with double-recrossing

trajectories is greater for q_{oLMax} than for r_{ion}, as shown in Fig. 4.8. We now examine the results from iLMax, which was designed specifically to give high transmission coefficients.

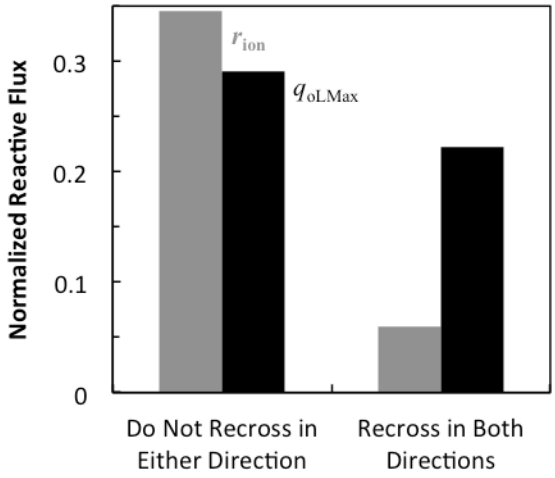

Figure 4.8. The normalized reactive flux for coordinates r_{ion} (gray) and q_{oLMax} (black) categorized by forward and time-reversed trajectories that do not recross or that both recross the dividing surface. The transmission coefficient κ is the difference between the two columns, e.g. $\kappa[r_{ion}] = 0.34 - 0.06 = 0.29$. The increased fraction of double recrossings in this ensemble is why q_{oLMax} gives a small transmission coefficient even though the committor distribution is narrower than for r_{ion} (Fig. **4.7**).

4.3.5 Inertial Likelihood Maximization (iLMax)

Including the velocity, acceleration, and jerk of r_{ion} increases $\Delta \ln L$ to 91. The reaction coordinate is further improved by adding solvent order parameters to r_{ion}, as shown in Table 4.2. The best among our trial coordinates, with $\Delta \ln L = 255$, is a function of r_{ion}, N_B, and the interionic water density ρ_{ii}. We label this coordinate q_{iLMax}. ρ_{ii} is calculated using a Gaussian-shaped indicator function centered at the midpoint between the ions. Fig. 4.7(b) and (c) shows the committor distributions from oLMax and iLMax are nearly identical. In contrast to the inconsistent results obtained from oLMax, Fig. 4.5 (dashed curve) shows the transmission coefficient $\kappa[q_{iLMax}]$ increases to 0.43 ± 0.02 consistent with coordinate improvement.

Table 4.2. Select Results from Inertial Likelihood Maximization[a]

$q_T(z)$[b]	c_A	c_V	$\Delta \ln L$[c]
$1.3\, r_{ion} - 4.8$	9.7	41	+91
$0.35\, r_{ion} + 75\, \rho_{ii} - 2.5$	-6.3	50	+233
$0.58\, r_{ion} + 0.18\, N_B + 53\, \rho_{ii} - 1.8$	-0.015	51	+255

[a]The best coordinate is indicated by the highest likelihood.
[b]The reaction coordinate is $q(\mathbf{z}) = q_T(\mathbf{z}) + c_A\, q_T(\mathbf{z})$. Units are: r_{ion} (Å), ρ_{ii} (Å^{-3}), N_B (unitless).
[c]Likelihoods $\Delta \ln L$ are reported in δL_{min} increments relative to that for r_{ion} from original likelihood maximization.

4.4 Discussion

Table 4.3 summarizes the transmission coefficients and committor distribution variances for the coordinates identified in this work. For q_{iLMax}, the standard deviation σ of the distribution of committors is 0.24. Clearly, there is room to further improve the description of the committor probability. Ballard and Dellago [69] showed that detailed nonlocal information from the first three solvation shells is needed to accurately predict the committor, whereas q_{iLMax} contains information only for water molecules in the first solvation shell. Interestingly, the coordinate p_B itself, by definition having $\sigma \approx 0$, gives $\kappa[p_B] = \kappa[q_{iLMax}]$ within the error bars on $\kappa[p_B]$. This result thus suggests that it is not always necessary to perfectly describe the committor to obtain an accurate dynamically corrected rate constant.

To examine the mechanism suggested by the reaction coordinate $q_{iLMax} = q(r_{ion}, \rho_{ii}, N_B)$, we compare the solvent components ρ_{ii} and N_B. Fig. 4.9 shows two transition state configurations with the same interionic water density, $\rho_{ii} = 0.016$ Å^{-3}, but unequal N_B. In Fig. 4.9(a), the 3 interionic waters are oriented randomly and $N_B \approx 0$. By contrast, in Fig. 4.9(b), $N_B = 1$ because the lone interionic water is oriented with a hydrogen toward Cl^- and the oxygen toward Na^+. The ions in Fig. 4.9(a), despite being farther apart than in 4.9(b) and

having similar interionic water density, are just as likely to recombine because the interionic waters are orientationally disordered. All waters proximal to the ion-pair midpoint contribute to ρ_{ii}, whereas a water molecule must be appropriately oriented to contribute to N_B. Of coordinates with 2 components, $\Delta \ln L[q(r_{ion}, \rho_{ii})] = 233$ while $\Delta \ln L[q(r_{ion}, N_B)] = 218$, showing that the presence of interionic waters is of primary importance to the reaction mechanism while the orientation of those water molecules is of secondary importance.

Table 4.3. Dividing Surface Properties for Different Reaction Coordinates

Reaction Coordinate	$q_{‡}$[a]	κ[b]	σ[c]
$q_{iLMax} = q(r_{ion}, \rho_{ii,} N_B)$	0	0.43	0.24
p_B	0.5	0.38	0.02
r_{ion}	3.7 Å	0.29	0.31
$q_{LMax} = q(r_{ion}, N_B, r_{opt})$	0	0.07	0.23

[a] $q_{‡}$ denotes the value of the reaction coordinate at the dividing surface.
[b] The transmission coefficient κ is 1 when TST is exact and has lower bound of 0.
[c] The standard deviation σ of the committor distribution is 0 for a perfect coordinate and at worst 0.5.

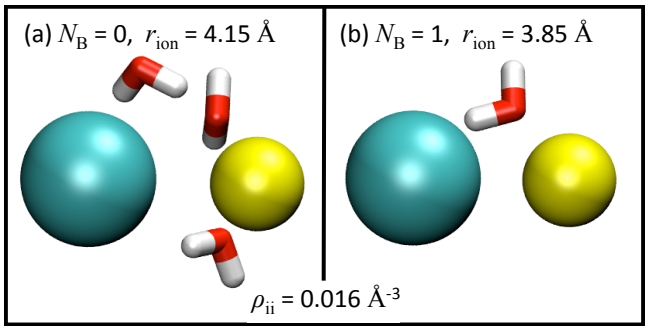

Figure 4.9. Different configurations from the transition state ensemble $q_{iLMax} = q(r_{ion}, \rho_{ii}, N_B) = 0$ showing Cl^- (blue), Na^+ (yellow) and interionic waters. For both configurations $\rho_{ii} = 0.016$ Å^{-3}, but in **(a)** $N_B \approx 0$ because all waters are randomly oriented, whereas in **(b)** $N_B = 1$ because the water is oriented with a hydrogen toward Cl^- and the oxygen toward Na^+.

By projecting the free energy onto the components of $q(r_{ion}, \rho_{ii}, N_B)$, we gain additional insight. Fig. 4.10(a) shows a surface similar to the harmonic model suggested by Truhlar and Garrett. On the other hand, Fig. 4.10(b) shows a rough, anharmonic surface with three distinct and parallel channels. The channels are parallel to r_{ion} and correspond to 0, 1 or 2 waters that bridge the ion pair. Within each channel, the transition states are at different locations along r_{ion}. Ion pairs with $N_B = 0$, 1, and 2 bridging waters must be approximately at $r_{ion} = 4.2$, 3.9 and 3.6 Å apart to be transition states, respectively. There is some scatter around these distances because the ρ_{ii} coordinate also modulates the location of the transition state.

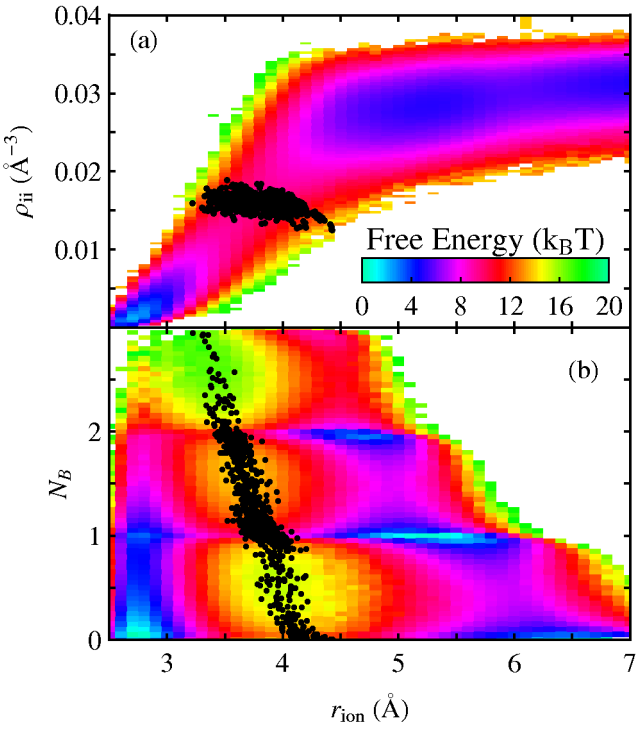

Figure 4.10. Free energy projected onto coordinates **(a)** r_{ion} and ρ_{ii}, **(b)** r_{ion} and N_B. Transition state $q_{iLMax} = q(r_{ion}, \rho_{ii}, N_B) = 0$ configurations are projected onto each surface as black points. In **(a)**, the surface resembles a harmonic valley. In **(b)**, the surface reveals separate transition states between intermediates along multiple channels.

The free energy surface $F(r_{ion}, \rho_{ii})$ looks like a harmonic valley and might be well described by bilinearly coupling ρ_{ii} to r_{ion}. However, the free energy surface $F(r_{ion}, N_B)$ reveals multiple channels and shallow intermediates in the ion-pair dissociation process. After crossing through the transition state, a trajectory reaching one of the shallow intermediates (e.g. $N_B = 2$, $r_{ion} = 4.7$ Å) may continue down the cascade to $N_B = 1$ and $N_B = 0$, or the ions may recombine at $N_B = 2$. Note that recrossing induced by shallow intermediates along the reaction pathway has also been reported for gas phase S_N2 reactions.[77-79] While it may be possible in full phase space,[80-82] no simple rotation of the reaction coordinate in configuration space will remove the shallow intermediates in the free energy landscape or the nonlinear forces from anharmonicity in the potential energy surface. We conclude that these features make recrossing an intrinsic part of the ion-pair dissociation dynamics.

4.5 Conclusions

We investigated dynamics of the committor for NaCl dissociation in TIP3P water. Non-reactive trajectories cross (and recross) the $p_B = ½$ isosurface giving $\kappa[p_B] \approx 0.38$. Consequently, as we have shown, a perfect reaction coordinate or dividing surface with no-recrossing cannot exist for ion-pair dissociation. Although recrossing cannot be entirely eliminated, inertial likelihood maximization did find a solvent coordinate that improves upon the ion-pair distance coordinate r_{ion}. This reaction coordinate combines r_{ion} with the density of water molecules between the two ions, and the number of waters jointly coordinated to them. The new coordinate gave a committor distribution with a broad peak near $p_B = ½$ and $\kappa = 0.43$. By comparison the ion-pair distance coordinate gives $\kappa[r_{ion}] = 0.29$ and a bimodal committor distribution.

The improvement in κ confirms the dynamical importance of specific solvent coordinates that were identified by inertial likelihood maximization. In contrast to a harmonic valley landscape of VTST approximations, the free energy landscape involving the coordinates identified by inertial likelihood maximization reveals multiple channels, shallow intermediates, and strong anharmonicity. These features of the free energy landscape correspond to physical details of the mechanism that make recrossing unavoidable for NaCl dissociation in water.

4.6 References

1. Reichardt, C.; Welton, T.: *Solvents and Solvent Effects in Organic Chemistry*; Wiley-VCH: Weinheim, Germany, 2011.
2. Voth, G. A.; Hochstrasser, R. M. "Transition State Dynamics and Relaxation Processes in Solutions: A Frontier of Physical Chemistry." *J. Phys. Chem.* **1996**, *100*, 13034-13049.
3. Castejon, H.; Wiberg, K. B. "Solvent effects on methyl transfer reactions. 1. The Menshutkin Reaction." *J. Am. Chem. Soc.* **1999**, *121*, 2139-2146.
4. Ensing, B.; Meijer, E. J.; Blöchl, P. E.; Baerends, E. J. "Solvation Effects on the SN2 Reaction between CH3Cl and Cl- in Water." *J. Phys. Chem. A* **2001**, *105*, 3300-3310.
5. Kramers, H. A. "Brownian motion in field of force and diffusion model of chemical reactions." *Physica* **1940**, *7*, 284-304.
6. Grote, R. F.; Hynes, J. T. "The Stable States Picture of Chemical-Reactions .2. Rate Constants for Condensed and Gas-Phase Reaction Models." *J. Chem. Phys.* **1980**, *73*, 2715-2732.
7. Berezhkovskii, A. M.; Pollak, E.; Zitserman, V. Y. "Activated Rate-Processes - Generalization of the Kramers-Grote-Hynes and Langer Theories." *J. Chem. Phys.* **1992**, *97*, 2422-2437.
8. Berne, B. J.; Borkovec, M.; Straub, J. E. "Classical and Modern Methods in Reaction-Rate Theory." *J. Phys. Chem.* **1988**, *92*, 3711-3725.
9. Anna, J. M.; Kubarych, K. J. "Watching solvent friction impede ultrafast barrier crossings: A direct test of Kramers theory." *J. Chem. Phys.* **2010**, *133*, 174506.
10. Eyring, H. "Activated complex in chemical reactions." *J. Chem. Phys.* **1935**, *3*, 107-115.
11. Miller, W. H. "Importance of Nonseparability in Quantum-Mechanical Transition-State Theory." *Acc. Chem. Res.* **1976**, *9*, 306-312.
12. Wigner, E. "Calculation of the Rate of Elementary Association Reactions." *J. Chem. Phys.* **1937**, *5*, 720-725.
13. Eyring, H. "Transmission Coefficient in Reaction Rate." *Rev. Mod. Phys.* **1962**, *34*, 616.
14. Keck, J. C. "Variational Theory of Chemical Reaction Rates Applied to Three-Body Recombinations." *J. Chem. Phys.* **1960**, *32*, 1035-1050.
15. Chandler, D. "Statistical mechanics of isomerization dynamics in liquids and the transition state approximation." *J. Chem. Phys.* **1978**, *68*, 2959-2970.

16. Truhlar, D. G.; Gao, J. L.; Alhambra, C.; Garcia-Viloca, M.; Corchado, J.; Sanchez, M. L.; Villa, J. "The incorporation of quantum effects in enzyme kinetics modeling." *Acc. Chem. Res.* **2002**, *35*, 341-349.
17. Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. "How enzymes work: Analysis by modern rate theory and computer simulations." *Science* **2004**, *303*, 186-195.
18. Hammes-Schiffer, S. "Hydrogen tunneling and protein motion in enzyme reactions." *Acc. Chem. Res.* **2006**, *39*, 93-100.
19. Peters, B.; Trout, B. L. "Obtaining reaction coordinates by likelihood maximization." *J. Chem. Phys.* **2006**, *125*, 054108.
20. van Erp, T. S. "Efficiency analysis of reaction rate calculation methods using analytical models I: The two-dimensional sharp barrier." *J. Chem. Phys.* **2006**, *125*, 174106.
21. van der Zwan, G.; Hynes, J. T. "Nonequilibrium solvation dynamics in solution reactions." *J. Chem. Phys.* **1983**, *78*, 4174-4185.
22. Tully, J. C. "Chemical dynamics at metal surfaces." *Annu. Rev. Phys. Chem.* **2000**, *51*, 153-178.
23. Truhlar, D. G.; Garrett, B. C. "Variational transition-state theory." *Acc. Chem. Res.* **1980**, *13*, 440-448.
24. Pollak, E. "Variational Transition-State Theory for Reactions in Condensed Phases." *J. Chem. Phys.* **1991**, *95*, 533-539.
25. Vanden-Eijnden, E.; Tal, F. A. "Transition state theory: Variational formulation, dynamical corrections, and error estimates." *J. Chem. Phys.* **2005**, *123*, 184103.
26. Hanggi, P.; Talkner, P.; Borkovec, M. "Reaction-rate theory: fifty years after Kramers." *Rev. Mod. Phys.* **1990**, *62*, 251.
27. Laidler, K. J.: *Chemical Kinetics*; Harper & Row: Cambridge, 1987.
28. Jensen, F.: *Introduction to computational chemistry*; Wiley: West Sussex, 1999.
29. Ariel, G.; Vanden-Eijnden, E. "Testing transition state theory on Kac-Zwanzig model." *J. Stat. Phys.* **2007**, *126*, 43-73.
30. Peters, B. "Inertial likelihood maximization for reaction coordinates with high transmission coefficients." *Chem. Phys. Lett.* **2012**, *554*, 248-253.
31. Rai, S. N.; Truhlar, D. G. "Variational Transition-State Theory Calculations for an Atom-Radical Reaction with No Saddle-Point - O+OH." *J. Chem. Phys.* **1983**, *79*, 6046-6059.
32. Hu, H.; Kobrak, M. N.; Xu, C. S.; Hammes-Schiffer, S. "Reaction path Hamiltonian analysis of dynamical solvent effects for a Claisen rearrangement and a Diels-Alder reaction." *J. Phys. Chem. A* **2000**, *104*, 8058-8066.
33. Torrie, G. M.; Valleau, J. P. "Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling." *J. Comput. Phys.* **1977**, *23*, 187-199.
34. Iannuzzi, M.; Laio, A.; Parrinello, M. "Efficient exploration of reactive potential energy surfaces using Car-Parrinello molecular dynamics." *Phys. Rev. Lett.* **2003**, *90*, 238302.
35. Truhlar, D. G.; Hase, W. L.; Hynes, J. T. "Current status of transition-state theory." *J. Phys. Chem.* **1983**, *87*, 2664-2682.
36. Truhlar, D. G.; Schenter, G. K.; Garrett, B. C. "Inclusion of nonequilibrium continuum solvation effects in variational transition state theory." *J. Chem. Phys.* **1993**, *98*, 5756.
37. Schenter, G. K.; Garrett, B. C.; Truhlar, D. G. "The role of collective solvent coordinates and nonequilibrium solvation in charge-transfer reactions." *J. Phys. Chem. B* **2001**, *105*, 9672-9685.

38. Peters, B.; Beckham, G. T.; Trout, B. L. "Extensions to the likelihood maximization approach for finding reaction coordinates." *J. Chem. Phys.* **2007**, *127*, 034109.
39. Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. "Transition path sampling: Throwing ropes over rough mountain passes, in the dark." *Annu. Rev. Phys. Chem.* **2002**, *53*, 291-318.
40. Du, R.; Pande, V. S.; Grosberg, A. Y.; Tanaka, T.; Shakhnovich, E. S. "On the transition coordinate for protein folding." *J. Chem. Phys.* **1998**, *108*, 334-350.
41. Geissler, P. L.; Dellago, C.; Chandler, D. "Kinetic pathways of ion pair dissociation in water." *J. Phys. Chem. B* **1999**, *103*, 3706-3710.
42. Rhee, Y. M.; Pande, V. S. "One-dimensional reaction coordinate and the corresponding potential of mean force from commitment probability distribution." *J. Phys. Chem. B* **2005**, *109*, 6780-6786.
43. Berezhkovskii, A.; Szabo, A. "One-dimensional reaction coordinates for diffusive activated rate processes in many dimensions." *J. Chem. Phys.* **2005**, *122*, 014503.
44. Ren, W.; Vanden-Eijnden, E. "Transition pathways in complex systems: Reaction coordinates, isocommittor surfaces, and transition tubes." *Chem. Phys. Lett.* **2005**, *413*, 242-247.
45. E, W.; Vanden-Eijnden, E. "Towards a theory of transition paths." *J. Stat. Phys.* **2006**, *123*, 503-523.
46. Beckham, G. T.; Peters, B. "Optimizing Nucleus Size Metrics for Liquid-Solid Nucleation from Transition Paths of Near-Nanosecond Duration." *J. Phys. Chem. Lett.* **2011**, *2*, 1133-1138.
47. Peters, B. "Using the histogram test to quantify reaction coordinate error." *J. Chem. Phys.* **2006**, *125*, 241101.
48. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. "Comparison of Simple Potential Functions For Simulating Liquid Water." *J. Chem. Phys.* **1983**, *79*, 926-935.
49. Fennell, C. J.; Bizjak, A.; Vlachy, V.; Dill, K. A. "Ion Pairing in Molecular Simulations of Aqueous Alkali Halide Solutions." *J. Phys. Chem. B* **2009**, *113*, 6782-6791.
50. Darden, T.; York, D.; Pedersen, L. "Particle Mesh Ewald - An N.Log(N) Method for Ewald Sums in Large Systems." *J. Chem. Phys.* **1993**, *98*, 10089-10092.
51. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. "Scalable molecular dynamics with NAMD." *J. Comput. Chem.* **2005**, *26*, 1781-1802.
52. Hoover, W. G. "Constant-Pressure Equations of Motion." *Phys. Rev. A* **1986**, *34*, 2499-2500.
53. Ferrenberg, A. M.; Swendsen, R. H. "Optimized Monte-Carlo Data-Analysis." *Phys. Rev. Lett.* **1989**, *63*, 1195-1198.
54. Peters, B. "Recent advances in transition path sampling: accurate reaction coordinates, likelihood maximisation and diffusive barrier-crossing dynamics." *Mol. Simul.* **2010**, *36*, 1265-1281.
55. Karim, O. A.; McCammon, J. A. "Dynamics of a Sodium-Chloride Ion-Pair in Water." *J. Am. Chem. Soc.* **1986**, *108*, 1762-1766.
56. Ciccotti, G.; Ferrario, M.; Hynes, J. T.; Kapral, R. "Dynamics of Ion-Pair Interconversion in a Polar-Solvent." *J. Chem. Phys.* **1990**, *93*, 7137-7147.

57. Rey, R.; Guardia, E. "Dynamic Aspects of the Na+-Cl- Ion-Pair Association in Water." *J. Phys. Chem.* **1992**, *96*, 4712-4718.
58. Dang, L. X.; Smith, D. E. "Molecular-Dynamics Simulations of Aqueous Ionic Clusters Using Polarizable Water." *J. Chem. Phys.* **1993**, *99*, 6950-6956.
59. Truhlar, D. G.; Garrett, B. C. "Multidimensional transition state theory and the validity of Grote-Hynes theory." *J. Phys. Chem. B* **2000**, *104*, 1069-1072.
60. Onsager, L. "Initial Recombination of Ions." *Physical Review* **1938**, *54*, 554-557.
61. Piana, S.; Jones, F.; Gale, J. D. "Assisted desolvation as a key kinetic step for crystal growth." *J. Am. Chem. Soc.* **2006**, *128*, 13568-13574.
62. Ruiz-Agudo, E.; Urosevic, M.; Putnis, C. V.; Rodriguez-Navarro, C.; Cardell, C.; Putnis, A. "Ion-specific effects on the kinetics of mineral dissolution." *Chem. Geol.* **2011**, *281*, 364-371.
63. Ma, A.; Dinner, A. R. "Automatic method for identifying reaction coordinates in complex systems." *J. Phys. Chem. B* **2005**, *109*, 6769-6779.
64. Bolhuis, P. G.; Dellago, C.; Chandler, D. "Reaction coordinates of biomolecular isomerization." *Proc. Natl. Acad. Sci.* **2000**, *97*, 5877-5882.
65. Hush, N. S. "Homogeneous and heterogeneous optical and thermal electron transfer." *Electrochim. Acta* **1968**, *13*, 1005-1023.
66. Mulliken, R. S. "Molecular Compounds and their Spectra. II." *J. Am. Chem. Soc.* **1952**, *74*, 811-824.
67. Marcus, R. A. "Generalization of the Activated Complex Theory of Reaction Rates. II. Classical Mechanical Treatment." *J. Chem. Phys.* **1964**, *41*, 2624-2633.
68. Hynes, J. T.: Crossing the transition state in solution. In *Solvent Effects and Chemical Reactivity*; Springer Netherlands: Dordrecht, Netherlands, 2002; pp 231-258.
69. Ballard, A. J.; Dellago, C. "Toward the Mechanism of Ionic Dissociation in Water." *J. Phys. Chem. B* **2012**, *116*, 13490-13497.
70. Marti, J.; Csajka, F. S. "The aqueous solvation of sodium chloride: A Monte Carlo transition path sampling study." *J. Chem. Phys.* **2000**, *113*, 1154-1161.
71. van der Zwan, G.; Hynes, J. T. "Dynamical Polar-Solvent Effects on Solution Reactions - A Simple Continuum Model." *J. Chem. Phys.* **1982**, *76*, 2993-3001.
72. Warshel, A. "Dynamics of Reactions in Polar-Solvents - Semi-classical Trajectory Studies of Electron-Transfer and Proton-Transfer Reactions." *J. Phys. Chem.* **1982**, *86*, 2218-2224.
73. Chau, P. L.; Hardwick, A. J. "A new order parameter for tetrahedral configurations." *Mol. Phys.* **1998**, *93*, 511-518.
74. DeMille, R. C.; Molinero, V. "Coarse-grained ions without charges: Reproducing the solvation structure of NaCl in water using short-ranged potentials." *J. Chem. Phys.* **2009**, *131*, 034107.
75. Gertner, B. J.; Wilson, K. R.; Hynes, J. T. "Nonequilibrium Solvation Effects on Reaction-Rates for Model SN2 Reactions in Water." *J. Chem. Phys.* **1989**, *90*, 3537-3558.
76. Kohen, D.; Tannor, D. J. "Phase-Space Distribution Function Formulation of the Method of Reactive Flux - Memory Friction." *J. Chem. Phys.* **1995**, *103*, 6013-6020.
77. Cho, Y. J.; Vande Linde, S. R.; Zhu, L.; Hase, W. L. "Trajectory studies of SN2 nucleophilic substitution. II. Nonstatistical central barrier recrossing in the Cl- + CH3Cl system." *J. Chem. Phys.* **1992**, *96*, 8275-8287.

78. Hase, W. L. "Simulations of Gas-Phase Chemical Reactions: Applications to SN2 Nucleophilic Substitution." *Science* **1994**, *266*, 998-1002.
79. Sun, L.; Hase, W. L.; Song, K. "Trajectory Studies of SN2 Nucleophilic Substitution. 8. Central Barrier Dynamics for Gas Phase Cl- + CH3Cl." *J. Am. Chem. Soc.* **2001**, *123*, 5753-5756.
80. Hernandez, R.; Uzer, T.; Bartsch, T. "Transition state theory in liquids beyond planar dividing surfaces." *Chem. Phys.* **2010**, *370*, 270-276.
81. Uzer, T.; Jaffe, C.; Palacian, J.; Yanguas, P.; Wiggins, S. "The geometry of reaction dynamics." *Nonlinearity* **2002**, *15*, 957.
82. Komatsuzaki, T.; Berry, R. "Local regularity and non-recrossing path in transition state—a new strategy in chemical reaction theories." *J. Mol. Struct. Theochem* **2000**, *506*, 55-70.

Chapter 5

An Existence Test for Dividing Surfaces Without Recrossing

A central goal of reaction rate theory is to obtain simple dynamical models which accurately reflect the mechanism of complex barrier crossing processes. The simplest model is transition state theory (TST) [1]. TST assumes (1) a dividing surface that separates reactants from products, (2) an equilibrium population of states on the dividing surface, and (3) a simple dynamics in which all trajectories crossing the dividing surface from reactant to product continue on to the product state without ever recrossing. If trajectories do recross the dividing surface, then the TST rate constant k_{TST} will overestimate the true rate constant k [2]. Time-dependent coordinate systems [3] and reaction coordinates that explicitly include momentum space degrees of freedom [4,5] can help eliminate recrossing, but TST is most easily implemented for static coordinate systems and dividing surfaces in configuration space. Therefore, our discussion is restricted to dividing surfaces of the type $q(\mathbf{x}) = q_{\ddagger}$, where $q(\mathbf{x})$ is a collective variable computed from atomic configurational coordinates $\mathbf{x}$ and $q_{\ddagger}$ is the value of the collective variable at the dividing surface. The variational transition state theory (VTST) seeks dividing surfaces with minimal recrossing to make TST rate estimates as accurate as possible [6-11].

When the dynamical bottleneck is a high saddle on the potential energy surface, a dividing surface with very few recrossings can be constructed based on a harmonic approximation to the potential energy surface [12,13]. The nearly perfect dividing surface passes through the saddle point and is oriented perpendicular to the unstable vibrational mode in mass-weighted coordinates [14,15]. The resulting multidimensional harmonic TST is widely used in analyzing chemical reactions.

When the saddle has strong anharmonic features at the considered temperature or when transition paths can take numerous routes on a rugged energy landscape, a full configuration space optimization of $q(\mathbf{x}) = q_{\ddagger}$ is not feasible. In these cases the true coordinate independent rate constant k can be computed using an approximate dividing surface and reaction coordinate q as

$$k = \kappa[q]\, k_{TST}[q] \tag{5.1}$$

where the transmission coefficient κ[q] is a correction for dynamical recrossing errors in k_{TST}[q]. κ can be obtained numerically from the reactive flux correlation functions using the full multidimensional dynamics. Alternatively, the one-dimensional models of Kramers theory [16] and Grote-Hynes theory [17] (GHT) invoke friction in the reaction coordinate dynamics to estimate the recrossing corrections. Kramers theory utilizes Markovian friction that implies the solvent instantly equilibrates to the reacting solute. GHT generalizes Kramers theory to account for more realistic solvent response times and non-Markovian friction. GHT has successfully been used to compute rates for solvated isomerizations [18,19], S_N2 reactions [20-22], dissociation reactions [23-28], and enzymatic catalysis [29,30].

GHT begins from a generalized Langevin equation (GLE) for the motion of a mass-weighted coordinate x

$$\ddot{x} = \omega^2 x - \int_0^t d\tau\, \eta(t-\tau)\, \dot{x}(\tau) + R, \tag{5.2}$$

where ω is the unstable frequency associated with a harmonic approximation to the potential of mean force $V(x)$ at the barrier top, R is a random force and η is a friction memory kernel which obeys the fluctuation-dissipation theorem

$$k_B T\, \eta(t) = \langle R(0) R(t) \rangle \tag{5.3}$$

The brackets indicate a thermal average. In Eq. (5.2) the time-dependence of x and R has been omitted except where correlating quantities at two different times. The rate constant from GHT is the product of k_{TST} and a transmission coefficient κ_{GH}

$$\kappa_{GH} = \frac{\lambda}{\omega} \tag{5.4}$$

where the reactive frequency λ is the largest, positive root to the GH equation

$$\lambda^2 + \lambda\, \tilde{\eta}(\lambda) - \omega^2 = 0 \tag{5.5}$$

$\tilde{\eta}(\lambda)$ is the frequency-dependent Laplace transform of the time-dependent friction $\eta(t)$. In practice computing κ_{GH} only requires $\eta(t)$ at $x = x_{\ddagger}$, which is evaluated from Eq. (5.3). It is important to note that GHT, in theory and in practice, does not begin with an optimized reaction coordinate. The k_{TST} contribution changes with the choice of x, so if GHT is to give correct overall rates, $\eta(t)$ and κ_{GH} must also change with different choices of x.

An important connection between multidimensional TST and GHT is often illustrated with a harmonic saddle model where friction emerges from a bath of bilinearly coupled harmonic oscillator (BCHO) modes [12,18,19,31-37]. The bath modes are combined with the potential of mean force $V(x)$ to form a multidimensional Hamiltonian

$$H = \frac{1}{2}p^2 + V(x) + \frac{1}{2}\sum_i \left[\dot{y}_i^2 + \omega_i^2\left(y_i + \frac{c_i}{\omega_i^2}x\right)^2\right] \tag{5.6}$$

which again starts from mass-weighted variables. p is the conjugate momentum to x, and y_i is the i^{th} bath mode with frequency ω_i and coupling constant c_i. The Hamiltonian dynamics along x from Eq. (5.6) reduce to a GLE [31], with the friction given by the cosine series

$$\eta(t) = \sum_i \frac{c_i^2}{\omega_i^2}\cos(\omega_i t) \tag{5.7}$$

When $V(x)$ is again replaced with its harmonic approximation, $V(x) \approx V(x_‡) - \omega^2 x^2/2$, then Eq. (5.6) becomes equivalent to Eq. (5.2) and, moreover, Eq. (5.6) becomes exactly solvable by multidimensional TST [18,19,32,33]. In the exact harmonic solution, the true reaction coordinate q—the unstable mode in the multidimensional space—mixes x with bath degrees of freedom. Recrossing is completely eliminated and the exact rate is obtained from TST without the transmission coefficient correction, as shown in Fig. 5.1a.

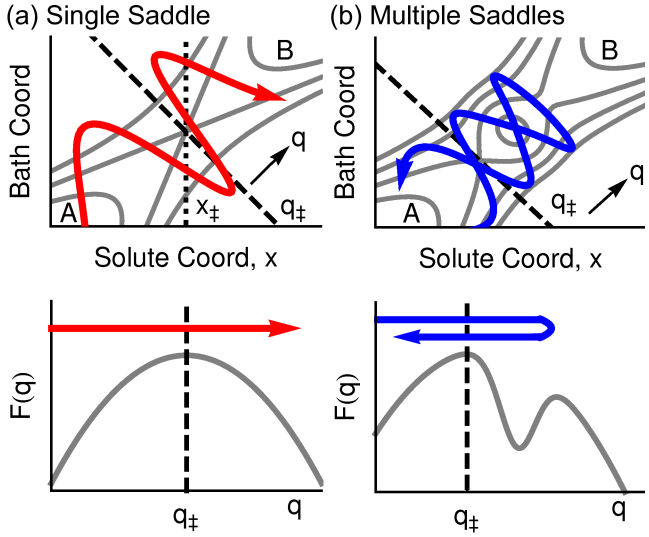

Figure 5.1. (Schematic) The upper panels depict deterministic trajectories on a two-dimensional model potential. In the lower panels, the trajectory and free energy F are projected onto the VTST coordinate q. **(a)** For a single, harmonic saddle, c.f. Eq. **(5.6)**, the $x = x_‡$ surface has recrossing, but the $q = q_‡$ surface does not. **(b)** An intermediate can cause trajectories to recross even the optimal dividing surface.

In the case of the BCHO model, the GHT calculation effectively corrects for the choice of a non-optimal dividing surface. In fact, the demonstration that a BCHO model is exactly equivalent to GHT and exactly solvable by VTST has been interpreted by many to mean that a successful application of GHT implies an underlying multidimensional BCHO model for which VTST would be exact. For example, Chandler has written that "GHT is a version of multidimensional transition state theory" [38]. Pollak commented “The equilibrium interaction of the bath with the system causes a change in the optimal transition state. Once this transition state is identified, there are no further corrections, dynamical or otherwise” [9]. Truhlar and Garrett wrote “the success of GHT … means that VTST with a suitable choice of solvent coordinate and an optimization of the orientation of the dividing surface does remove the recrossing” [39]. This perspective is supported by studies of the S_N2 reaction $Cl^- + CH_3Cl$ in water, for which κ_{GH} was computed from a GLE, Eqs. (5.2) and (5.3) [20-22], and identical κ_{GH} were later computed from a BCHO, Eqs. (5.6) and (5.7) [40]. The 'reaction coordinate' of GHT is defined using only solute degrees of freedom and any mechanistic role of the solvent is lumped into the dynamical transmission coefficient κ_{GH}. Therefore, the idea that VTST could specifically identify the mechanistic role of the solvent while also providing dynamically accurate rates is appealing.

However, Hynes argued that GHT is more general than VTST [38], and that the BCHO model is perhaps not the only way to obtain the GLE [22]. Dakhnovskii and Ovchinnikov espoused a similar perspective, noting "[BCHO] oscillators should be considered nonphysical…. We can only determine their combination which is equal to the memory function" [32]. In atomistic systems, motion along the reaction coordinate under the PMF might be coupled to an anharmonic bath, a bath of inter-coupled oscillators, a bath possessing

local intermediate traps (see Fig. 5.1b), or even to a system with multiple non-equivalent pathways all affecting the same transformation. Unfortunately, such highly non-linear dynamics are not amenable to analytic projections, so there is no rigorous demonstration that the dynamics projected onto a coordinate would follow a second order GLE like that of Eq. (5.2). In several cases where GHT has been demonstrated to give accurate rates [23-30], it has not yet been possible to identify the solvent coordinates or optimize dividing surfaces, by VTST or by other means. It therefore remains difficult to determine whether recrossing emerges from a non-optimal reaction coordinate, or from some intrinsic friction related to non-linearities in the dynamics.

The true origin of recrossing, in our view, depends on the system being investigated and on the coordinates being used. In developing simple models of complex barrier crossing processes, it would be useful to have a systematic test to determine whether friction-induced recrossing of the dividing surface can be eliminated by some better coordinate, or whether we should accept recrossing as inevitable and move forward with approximate GLE models of the dynamics.

VTST can provide exact rates with no further corrections only if a recrossing-free dividing surface exists. The claim that a no-recrossing surface exists has potentially falsifiable consequences. Suppose that a dividing surface exists in configuration space that is not recrossed by any deterministic trajectories. By definition, all trajectories initiated in A that reach the dividing surface must continue on to B. An A→A trajectory cannot share a common configuration **x**—coordinates of all atoms identical—with a B→B trajectory. We refer to these A→**x**→A and B→**x**→B as recrossing pairs. The observation of a recrossing pair proves that a no-recrossing dividing surface cannot exist because all possible dividing

surfaces are recrossed by the A→**x**→A trajectory, the B→**x**→B trajectory, or both, as shown in Fig. 5.2.

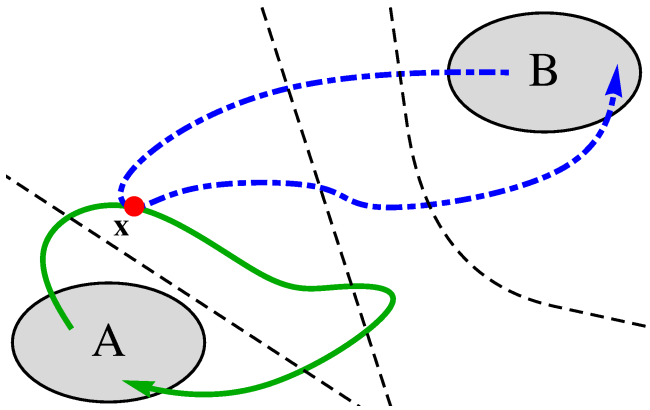

Figure 5.2. Schematic of a system with two stable states A and B for which a recrossing pair, an A→**x**→A trajectory (green) and a B→**x**→B trajectory (blue, dot-dashed), has been observed. Any possible dividing surface (e.g., black, dashed) is recrossed by the A→**x**→A trajectory, the B→**x**→B trajectory, or both.

Note that, formally speaking, even one recrossing pair proves the non-existence of a perfect dividing surface at all temperatures [12]. A recrossing pair, once observed at some temperature, may be more or less important in the complete path ensemble [41] at other temperatures, but it is a member of the complete path ensemble at all temperatures. In practice, however, trajectories with energies far above the activation energy have near-zero statistical weight and will have negligible effects on the transmission coefficient. Next, we discuss practical methods to sample the ensemble of trajectories at the temperature of interest.

Transition Path Sampling [42] (TPS) provides one way to identify recrossing-pairs. In the TPS algorithm a trial trajectory is generated from a randomly selected configuration, called a shooting point, on an existing transition path. If the trial trajectory is reactive, it is accepted and replaces the existing transition path. Each non-reactive trial trajectory is rejected in the sampling algorithm but could potentially be part of a recrossing-pair. Recrossing-pairs can be

identified by tracking the outcomes (e.g., A→A, B→B) for all non-reactive trajectories initiated from each shooting point configuration. Microcanonical (NVE) versions of TPS should be used to identify recrossing pairs so that the dynamics are not altered by the stochastic thermostat. In Example 1 below, we have used an NVE permutation shooting algorithm. In Example 3, we have used NPT Aimless Shooting [43] to generate an ensemble of shooting points with different volumes V, energies E, and initial momenta. However, these initial conditions were strictly used to propagate deterministic microcanonical trajectories for the detection of recrossing pairs.

5.1 Example 1: BCHO model

We first apply the test for recrossing-pairs to the BCHO model of Eq. (5.6). We anticipate that zero recrossing pairs will be observed because this model is proven to have a no-recrossing dividing surface [18,19,32,33]. Using an inverted parabola $V(x) = -\omega^2 x^2/2$ and 11 bath modes with the parameters in Ref. [44], we have harvested 100,000 trajectories by microcanonical TPS with a 50% acceptance ratio (i.e., 50% were transition paths and the balance were non-reactive trajectories). All sampled trajectories have the same total energy E = 11.5 $k_B T$ relative to the saddle point potential, where the energy is measured in units such that $k_B T = 1.0$. Multiple trajectories were initiated from 19,651 configurations. Not a single recrossing pair was observed from these data. For each configuration, there were either no non-reactive trajectories, all the non-reactive trajectories were A→A, or all B→B. In principle, the existence of a recrossing pair precludes a no-recrossing surface, but the absence of recrossing pair is not sufficient to prove existence of a perfect no-recrossing surface. In practice, when recrossing pairs are extremely rare a fully variational dividing surface optimization should lead to accurate TST rates.

5.2 Example 2: Diffusive dynamics

We next consider a system with purely diffusive dynamics, or equivalently the limiting Kramers regime of large friction. In this limit, $\kappa \to 0$ and a no-recrossing surface is not possible. The expected number of recrossing pairs to be observed with Aimless Shooting can be calculated analytically using the committor p_B. p_B gives the probability that a trajectory initiated from **x** will relax to B. [45,46] Since forward and backward trajectories from **x** are independent, the probability of generating a B→**x**→B trajectory is p_B^2 and an A→**x**→A is $(1-p_B)^2$. The fraction of recrossing pairs at each value of p_B must be

$$f(p_B) = 2p_B^2(1-p_B)^2 \tag{5.8}$$

f is sharply peaked near $p_B = ½$. Shooting points from Aimless Shooting are distributed with statistical weight [43]

$$p(p_B) \sim p_B(1-p_B) \tag{5.9}$$

From Eqs. (5.8) and (5.9), the expected number of recrossing pairs is

$$\langle f \rangle = \frac{\int_0^1 dp_B\, f(p_B) p(p_B)}{\int_0^1 dp_B\, p(p_B)} = \frac{8.6 \text{ recrossing pairs}}{100 \text{ configurations sampled}} \tag{5.10}$$

5.3 Example 3: Ion-pair dissociation

Finally, we consider an atomistic model of ion-pair dissociation in explicit solvent. Previous studies have confirmed the accuracy of GHT in predicting the rate of ion-pair dissociation [23,24]. It has been proposed, specifically in the context of this system, that a dividing surface optimized by VTST might eliminate the recrossing [39]. As yet neither an accurate reaction coordinate nor a dividing surface without recrossing has been found. Therefore, ion-pair dissociation provides an ideal system to test for the existence of a no-recrossing surface.

We performed 24,000 molecular dynamics trajectories of an OPLS [47] NaCl ion-pair in TIP3P [48] water. Trajectories were computed in NAMD [49] using deterministic microcanonical (NVE) dynamics and periodic boundary conditions. The shooting point configuration and momenta for each trajectory were taken from a previous [50] Aimless Shooting simulation performed at constant temperature T=300 K, pressure P=1 bar and number of water molecules N=394. The volume V and energy E of each shooting point are different, but are held constant when propagating the system forward or backward in time for the present study. The timestep was 0.25 fs and the electrostatic energy was computed using Particle Mesh Ewald [51].

Multiple trajectories were initiated from 4706 configurations and 419 recrossing pairs were observed. The ratio 8.9 recrossing pairs to 100 configurations sampled is similar in magnitude to Example 2, for which even a surface with few recrossings is impossible. The approximate agreement is perhaps coincidental, because Ballard and Dellago showed that ion-pair dissociation trajectories have significant inertial character [52]. Regardless, we can conclude that any trial dividing surface will have recrossings in a large portion of transition paths for this system. Clearly, the dynamical effects associated with ion-pair dissociation in explicit solvent cannot be eliminated by VTST. Because GHT provides an accurate transmission coefficient for ion-pair dissociation, this system provides a specific counterexample to the viewpoint that GHT and VTST have the same scope and underlying theoretical basis. Clearly GHT can (at least approximately) account for friction induced recrossings of non-linear origins that cannot be removed by a dividing surface optimization.

5.4 Conclusion

Many authors have asserted that an accurate rate constant from Grote-Hynes (GH) theory implies an underlying multidimensional system with a bilinearly coupled solvent for which VTST, if the appropriate solvent coordinates were known, would also become exact. This assertion has been difficult to support or refute with simulation data because of the extreme difficulties in optimizing dividing surfaces and reaction coordinates for reactions in solution. Here we note that VTST can only provide accurate rates—and therefore be equivalent to accurate GHT rates—when a surface with no recrossing exists. The assertion of an equivalence between GHT and VTST thus implies the existence of a no-recrossing surface, and the latter is potentially falsifiable. If a no-recrossing surface exists, then non-reactive trajectories containing an intermediate configuration **x** must all terminate as either reactants (A) or as products (B). A pair of A→**x**→A and B→**x**→B trajectories sharing a common point **x** is forbidden because any dividing surface would have to be crossed at least twice. We have studied two systems for which GHT is successful: a parabolic barrier bilinearly coupled to harmonic oscillators (BCHO) and atomistic ion-pair dissociation. Consistent with the well-known construction of a no-recrossing surface for the BCHO model, we observe no recrossing pairs. Transition path sampling data for ion-pair dissociation, however, reveals many recrossing pairs, proving that a no-recrossing surface (and even a surface with few recrossings) cannot exist. Because GHT does provide accurate transmission coefficients for ion-pair dissociation, our paper shows that GHT, contrary to some prior assertions, does have a broader scope of approximate validity than VTST.

5.5 References

1. Eyring, H. "Activated complex in chemical reactions." *J. Chem. Phys.* **1935**, *3*, 107-115.
2. Wigner, E. "Calculation of the Rate of Elementary Association Reactions." *J. Chem. Phys.* **1937**, *5*, 720-725.
3. Komatsuzaki, T.; Berry, R. "Local regularity and non-recrossing path in transition state—a new strategy in chemical reaction theories." *J. Mol. Struct. Theochem* **2000**, *506*, 55-70.
4. Uzer, T.; Jaffe, C.; Palacian, J.; Yanguas, P.; Wiggins, S. "The geometry of reaction dynamics." *Nonlinearity* **2002**, *15*, 957.
5. Hernandez, R.; Uzer, T.; Bartsch, T. "Transition state theory in liquids beyond planar dividing surfaces." *Chem. Phys.* **2010**, *370*, 270-276.
6. Keck, J. C. "Variational Theory of Chemical Reaction Rates Applied to Three-Body Recombinations." *J. Chem. Phys.* **1960**, *32*, 1035-1050.
7. Truhlar, D. G.; Garrett, B. C. "Variational transition-state theory." *Acc. Chem. Res.* **1980**, *13*, 440-448.
8. Truhlar, D. G.; Hase, W. L.; Hynes, J. T. "Current status of transition-state theory." *J. Phys. Chem.* **1983**, *87*, 2664-2682.
9. Pollak, E. "Variational Transition-State Theory for Reactions in Condensed Phases." *J. Chem. Phys.* **1991**, *95*, 533-539.
10. Vanden-Eijnden, E.; Tal, F. A. "Transition state theory: Variational formulation, dynamical corrections, and error estimates." *J. Chem. Phys.* **2005**, *123*, 184103.
11. van Erp, T. S. "Dynamical Rare Event Simulation Techniques for Equilibrium and Nonequilibrium Systems." *Adv. Chem. Phys.* **2012**, *151*, 27-60.
12. Hanggi, P.; Talkner, P.; Borkovec, M. "Reaction-rate theory: fifty years after Kramers." *Rev. Mod. Phys.* **1990**, *62*, 251.
13. Laidler, K. J.: *Chemical Kinetics*; Harper & Row: Cambridge, 1987.
14. Wilson, E. B.: *Molecular vibrations: the theory of infrared and Raman vibrational spectra*; Dover: New York, 1955.
15. Jensen, F.: *Introduction to computational chemistry*; Wiley: West Sussex, 1999.
16. Kramers, H. A. "Brownian motion in field of force and diffusion model of chemical reactions." *Physica* **1940**, *7*, 284-304.
17. Grote, R. F.; Hynes, J. T. "The Stable States Picture of Chemical-Reactions .2. Rate Constants for Condensed and Gas-Phase Reaction Models." *J. Chem. Phys.* **1980**, *73*, 2715-2732.
18. van der Zwan, G.; Hynes, J. T. "Nonequilibrium solvation dynamics in solution reactions." *J. Chem. Phys.* **1983**, *78*, 4174-4185.
19. van der Zwan, G.; Hynes, J. T. "A simple dipole isomerization model for non-equilibrium solvation dynamics in reactions in polar solvents." *Chem. Phys.* **1984**, *90*, 21-35.
20. Bergsma, J. P.; Gertner, B. J.; Wilson, K. R.; Hynes, J. T. "Molecular-Dynamics of a Model SN2 Reaction in Water." *J. Chem. Phys.* **1987**, *86*, 1356-1376.
21. Gertner, B. J.; Bergsma, J. P.; Wilson, K. R.; Lee, S. Y.; Hynes, J. T. "Nonadiabatic Solvation Model for SN2 Reactions in Polar-Solvents." *J. Chem. Phys.* **1987**, *86*, 1377-1386.
22. Gertner, B. J.; Wilson, K. R.; Hynes, J. T. "Nonequilibrium Solvation Effects on Reaction-Rates for Model SN2 Reactions in Water." *J. Chem. Phys.* **1989**, *90*, 3537-3558.

23. Ciccotti, G.; Ferrario, M.; Hynes, J. T.; Kapral, R. "Dynamics of Ion-Pair Interconversion in a Polar-Solvent." *J. Chem. Phys.* **1990**, *93*, 7137-7147.
24. Rey, R.; Guardia, E. "Dynamic Aspects of the Na+-Cl- Ion-Pair Association in Water." *J. Phys. Chem.* **1992**, *96*, 4712-4718.
25. Dang, L. X.; Smith, D. E. "Molecular-Dynamics Simulations of Aqueous Ionic Clusters Using Polarizable Water." *J. Chem. Phys.* **1993**, *99*, 6950-6956.
26. Dang, L. X.; Schenter, G. K.; Chang, T.-M.; Kathmann, S. M.; Autrey, T. "Role of Solvents on the Thermodynamics and Kinetics of Forming Frustrated Lewis Pairs." *J. Phys. Chem. Lett.* **2012**, *3*, 3312-3319.
27. Nguyen, P. T. M.; Nguyen, V. T.; Annapureddy, H. V. R.; Dang, L. X.; Do, D. D. "Thermodynamics and kinetics of Na+/K+-formate ion pairs association in polarizable water: A molecular dynamics study." *Chem. Phys. Lett.* **2012**, *554*, 90-95.
28. Annapureddy, H. V. R.; Dang, L. X. "Pairing Mechanism among Ionic Liquid Ions in Aqueous Solutions: A Molecular Dynamics Study." *J. Phys. Chem. B* **2013**, *117*, 8555-8560.
29. Roca, M.; Moliner, V.; Tunon, I.; Hynes, J. T. "Coupling between protein and reaction dynamics in enzymatic processes: Application of Grote-Hynes theory to catechol O-methyltransferase." *J. Am. Chem. Soc.* **2006**, *128*, 6186-6193.
30. Ruiz-Pernia, J. J.; Tunon, I.; Moliner, V.; Hynes, J. T.; Roca, M. "Dynamic effects on reaction rates in a Michael addition catalyzed by chalcone isomerase. Beyond the frozen environment approach." *J. Am. Chem. Soc.* **2008**, *130*, 7477-7488.
31. Zwanzig, R. "Nonlinear generalized Langevin equations." *J. Stat. Phys.* **1973**, *9*, 215-220.
32. Dakhnovskii, Y. I.; Ovchinnikov, A. A. "Decay of a metastable state: Comparison between the transition state theory and the generalized Kramers model." *Phys. Lett. A* **1985**, *113*, 147-150.
33. Pollak, E. "Theory of Activated Rate-Processes - a New Derivation of Kramers Expression." *J. Chem. Phys.* **1986**, *85*, 865-867.
34. Haynes, G. R.; Voth, G. A. "Reaction-Coordinate-Dependent Friction in Classical Activated Barrier Crossing Dynamics - When It Matters and When It Doesn't." *J. Chem. Phys.* **1995**, *103*, 10176-10182.
35. Voth, G. A.; Hochstrasser, R. M. "Transition State Dynamics and Relaxation Processes in Solutions: A Frontier of Physical Chemistry." *J. Phys. Chem.* **1996**, *100*, 13034-13049.
36. Schenter, G. K.; Garrett, B. C.; Truhlar, D. G. "The role of collective solvent coordinates and nonequilibrium solvation in charge-transfer reactions." *J. Phys. Chem. B* **2001**, *105*, 9672-9685.
37. Bao, J.-D.; Zhuo, Y.-Z.; Oliveira, F. A.; Hanggi, P. "Intermediate dynamics between Newton and Langevin." *Phys. Rev. E* **2006**, *74*, 061111.
38. Barbara, P.; Hynes, J. T.; Chandler, D.; Symons, M. C. R.; Abraham, M. H.; Wells, C. F.; Henchman, M.; Blandamer, M. J.; Bush, D.; Halawani, K. H.; Schroeder, J.; Troe, J.; Wolynes, P. G.; Suppan, P.; Friedman, H. L.; Hall, D. G. "General discussion." *Faraday Discuss.* **1988**, *85*, 341-364.
39. Truhlar, D. G.; Garrett, B. C. "Multidimensional transition state theory and the validity of Grote-Hynes theory." *J. Phys. Chem. B* **2000**, *104*, 1069-1072.
40. Gershinsky, G.; Pollak, E. "Variational Transition-State Theory - Application to a Symmetrical Exchange-Reaction in Water." *J. Chem. Phys.* **1995**, *103*, 8501-8512.

41. Rogal, J.; Bolhuis, P. G. "On the efficiency of biased sampling of the multiple state path ensemble." *J. Chem. Phys.* **2010**, *133*, 034101.
42. Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. "Transition path sampling: Throwing ropes over rough mountain passes, in the dark." *Annu. Rev. Phys. Chem.* **2002**, *53*, 291-318.
43. Peters, B.; Beckham, G. T.; Trout, B. L. "Extensions to the likelihood maximization approach for finding reaction coordinates." *J. Chem. Phys.* **2007**, *127*, 034109.
44. Peters, B. "Inertial likelihood maximization for reaction coordinates with high transmission coefficients." *Chem. Phys. Lett.* **2012**, *554*, 248-253.
45. Du, R.; Pande, V. S.; Grosberg, A. Y.; Tanaka, T.; Shakhnovich, E. S. "On the transition coordinate for protein folding." *J. Chem. Phys.* **1998**, *108*, 334-350.
46. Geissler, P. L.; Dellago, C.; Chandler, D. "Kinetic pathways of ion pair dissociation in water." *J. Phys. Chem. B* **1999**, *103*, 3706-3710.
47. Fennell, C. J.; Bizjak, A.; Vlachy, V.; Dill, K. A. "Ion Pairing in Molecular Simulations of Aqueous Alkali Halide Solutions." *J. Phys. Chem. B* **2009**, *113*, 6782-6791.
48. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. "Comparison of Simple Potential Functions For Simulating Liquid Water." *J. Chem. Phys.* **1983**, *79*, 926-935.
49. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. "Scalable molecular dynamics with NAMD." *J. Comput. Chem.* **2005**, *26*, 1781-1802.
50. Mullen, R. G.; Shea, J.-E.; Peters, B. "Transmission coefficients, committors, and solvent coordinates in ion-pair dissociation." *J. Chem. Theory Comput.* **2014**, *10*, 659-667.
51. Darden, T.; York, D.; Pedersen, L. "Particle Mesh Ewald - An N.Log(N) Method for Ewald Sums in Large Systems." *J. Chem. Phys.* **1993**, *98*, 10089-10092.
52. Ballard, A. J.; Dellago, C. "Toward the Mechanism of Ionic Dissociation in Water." *J. Phys. Chem. B* **2012**, *116*, 13490-13497.

Chapter 6

Conclusion

We have developed new, simple transition path sampling (TPS) methods that reduce the computational expense of simulating rare events. Permutation shooting rigorously preserves the total energy and momentum from one trajectory to the next making it a naturally suited to sampling microcanonical (NVE) dynamics. In contrast to existing methods, permutation shooting is easily adaptable to simulations with rigid water molecules. Like aimless shooting, permutation shooting only uses two potential shooting points per trajectory, regardless of the length of the transition paths. Therefore, fixed-length and flexible-length versions of aimless shooting and permutation shooting have identically simple Metropolis acceptance criteria. Additionally, flexible-length versions of TPS do not require the *a priori* choice of trajectory length and reduce the computational cost by terminating simulations once a trajectory enters a stable basin.

We applied permutation shooting and aimless shooting each to a different condensed phase system and analyzed the resulting data with likelihood maximization to identify the reaction coordinate. The location of a vacancy in a single-domain crystal diffuses as a series of particles hop into the vacant site. Each hop is an activated event that briefly deforms the local crystal structure. In Chapter 3, we compared two trial reaction coordinates that have

previously been used to describe vacancy hopping: position along the straight line path between donor and acceptor sites ξ and the bipolar distance q. We found that both ξ and q depend on the accuracy with which the donor and acceptor lattice sites are located. Likelihood maximization revealed that an energy minimization technique was not sufficiently accurate to compute the trial coordinate velocities. Instead, donor and acceptor sites were computed using a linear combination of neighboring particle positions. With this improved definition, q is the better coordinate, recrossing of the dividing surface is effectively eliminated and the rate can easily be computed from transition state theory (TST).

In Chapter 4, we investigated the role of water in ion-pair dissociation. Ion pairs do not dissociate in the gas phase, but salt dissolving in water is a common, everyday occurrence. Using aimless shooting and likelihood maximization, we identified two solvent mechanisms that influence ion-pair transition states. The primary mechanism occurs as a water molecule enters the region between the ions, as measured by the interionic water density. The secondary mechanism occurs as the water molecule orients to jointly coordinate both ions. The free energy projected onto the solvent coordinates reveals multiple channels, shallow intermediates, and strong anharmonicity. Consistent with such a complicated energy landscape, the best dividing surface in these coordinates does not eliminate recrossing and yields a transmission coefficient κ well below 1.

The ion-pair results led us to develop a test for the existence of a no-recrossing surface. This test uses unbiased dynamical trajectories without any dimensionality reduction, such as are generated using TPS. If a no-recrossing surfaces exists, then non-reactive trajectories containing an intermediate configuration $\mathbf{x}$ must all terminate as either reactants (A) or products (B). A pair of A→$\mathbf{x}$→A and B→$\mathbf{x}$→B trajectories sharing a common point $\mathbf{x}$ is

forbidden because any dividing surface would have to be crossed at least twice. TPS ion-pair dissociation data reveals many recrossing pairs, proving that a no-recrossing surface (and even a surface with few recrossings) cannot exist. Additionally, we developed a heuristic for estimating the maximum possible κ. An exact dividing surface has the highest possible κ and is composed entirely of $p_B = ½$ states. For ion-pair dissociation, κ of $p_B = ½$ states is approximately equal to κ of the optimized solvent coordinate dividing surface, further supporting the identified solvent mechanism.

Finally, the existence of a no-recrossing dividing surface has ramifications for the relationship between TST and theories that incorporate solvent friction, specifically Grote-Hynes theory (GHT). In the limit of a high, harmonic saddle crossing, TST and GHT give identical rates, leading some to assert that the two theories are also identical. However, TST can only provide accurate rates—and therefore be equivalent to accurate GHT rates—when a no-recrossing surface exists. Combining previous researchers results that GHT gives accurate rates for ion-pair dissociation, with our results that a no-recrossing surface can not exist, we see that GHT and TST are not equivalent.

Appendix A

Importance of Randomly Selecting Potential Shooting Points

In permutation shooting, each trajectory has two potential shooting points. One potential shooting point is the configuration used to generate the trajectory, $\mathbf{x}_0^o$. The other should be randomly selected from the forward (at $t = \Delta t$) or backward (at $t = -\Delta t$) half-trajectory to improve the algorithm's sampling.

To illustrate, consider a shooting point configuration $\mathbf{r}_0^o$ with a high probability of being on a transition path, $p(\mathrm{TP}|\mathbf{r}_0^o)$. Even large changes to the corresponding momentum $\mathbf{p}_0^o$ will likely yield a new transition path (see Fig. A.1A). The next accepted trajectory will rapidly diverge from the old trajectory, and the new potential shooting point $\mathbf{r}_{\Delta \mathrm{t}}^n$ will be dissimilar to $\mathbf{r}_{\Delta \mathrm{t}}^o$.

Now consider a trajectory for which $p(\mathrm{TP}|\mathbf{r})$ is low for both potential shooting points. For example, a shooting point pair near basin B will likely exhibit $p(\mathrm{TP}|\mathbf{r}_{\Delta \mathrm{t}}^o) \ll p(\mathrm{TP}|\mathbf{r}_0^o) \ll 1$ as shown in Fig. A.1B. In this circumstance, most new trajectories will be non-reactive. A new transition path, though rare, is most likely when shooting from $\mathbf{r}_0^o$, and even then only when the new momenta $\mathbf{p}_0^n$ are similar to the old $\mathbf{p}_0^o$, as in Fig. A.1C. The new path will diverge slowly from the old path and the timeslices $\mathbf{x}_0^o$ and $\mathbf{x}_{\Delta \mathrm{t}}^o$ will be similar to $\mathbf{x}_0^n$ and $\mathbf{x}_{\Delta \mathrm{t}}^n$,

respectively. If the alternate shooting point was always taken from the forward half-trajectory at $t = \Delta t$, the algorithm would effectively be limited to sampling trajectory space in the vicinity of the old path. Since the backward half-trajectory connects to the other basin, it likely goes through regions of high $p(\mathrm{TP}|\mathbf{r})$. If the alternate shooting point can come from the backward half-trajectory at $t = -\Delta t$, the shooting points can move away from a nearby basin even though the new transition path is similar to the old.

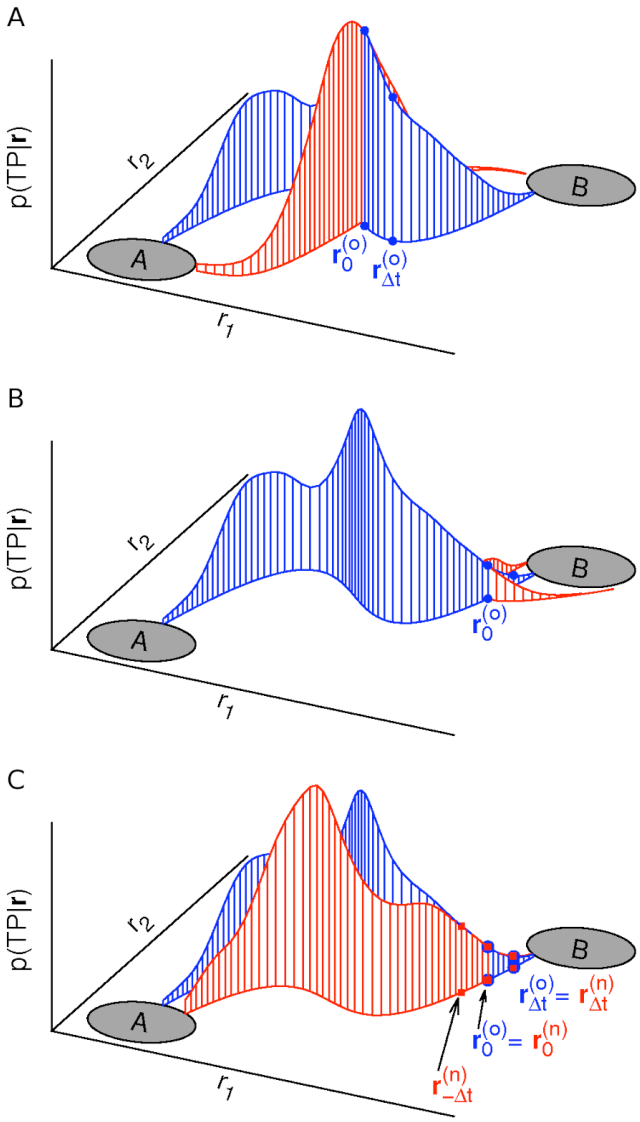

Figure A.1. Schematic showing the importance of selecting the new second shooting point randomly (see text). Configuration space is represented by two coordinates x_1 and x_2 and the z-axis shows the probability that configuration r is on a transition path, $p(\mathrm{TP}|\mathbf{r})$. The old trajectory is shown in blue. A new trajectory is shown in red. **(A)** The shooting points are near the middle of the transition region where $p(\mathrm{TP}|\mathbf{r})$ is highest. **(B, C)** The shooting points are near basin B where $p(\mathrm{TP}|\mathbf{r})$ is low.

Appendix B

Uncertainty in the Transmission Coefficient

In general, $\kappa[q]$ is the ratio of two conditional averages

$$\kappa = \frac{\langle\, \dot{q}\; \theta\big(q\big(t_{plateau}\big) - q_{\ddagger}\,\big\rangle_{\ddagger}}{\langle\, |\dot{q}|\,/\,2\rangle_{\ddagger}} \tag{B.1}$$

where q is the reaction coordinate, $q_{\ddagger}$ is the value of the reaction coordinate at the top of the barrier, θ is the Heaviside function, and t_{plateau} is a time greater than the molecular relaxation time but less than the reaction time [1]. The standard error of the mean for each average in Eq. (B.1) is given by the Central Limit Theorem

$$\sigma^2_{\langle x\rangle} = \frac{\langle x^2\rangle - \langle x\rangle^2}{n} \tag{B.2}$$

where n is the number of trajectories generated to calculate κ. Using the notation $\kappa = A\,/\,B$ for Eq. (B.1), the Taylor series expansion of κ gives

$$\frac{\delta\kappa}{\kappa} = \frac{\delta A}{A} - \frac{\delta B}{B} + O(\delta B^2 + \delta A \delta B) \tag{B.3}$$

Taking the square of Eq. (B.3) and then averaging gives an expression for σ_{κ}

$$\left(\frac{\sigma_\kappa}{\kappa}\right)^2 = \left(\frac{\sigma_A}{A}\right)^2 - \left(\frac{\sigma_B}{B}\right)^2 - 2\frac{Cov(A,B)}{A\,B} \tag{B.4}$$

Error bars in Fig. 4.5 were calculated from Eq. (B.4).

There is an additional source of uncertainty to consider when calculating $\kappa[p_B]$. The time derivative $\dot{p}_B$ necessary to compute $\kappa[p_B]$ will be distributed according to some probability $P(\dot{p}_B)$. The standard deviation $\sigma\dot{p}_B$ introduces uncertainty into averages calculated from a finite number of $\dot{p}_B$ samples. We approximate $\dot{p}_B$ by central finite difference from estimates of p_B (see Fig. 4.1). p_B is an intrinsic property of configuration **x**, while $\hat{p}_B$ is an estimate of p_B computed by initiating N trajectories from **x** and counting the fraction that go to B. The estimation uncertainty in $\hat{p}_B$ is

$$\delta\hat{p}_B = \sqrt{\frac{\hat{p}_B(1-\hat{p}_B)}{N}} \tag{B.5}$$

and the uncertainty in $\hat{\dot{p}}_B \approx \Delta\hat{p}_B/\Delta t$ is $\delta\hat{\dot{p}}_B = \sqrt{2}\,\delta\hat{p}_B / \Delta t$ (Fig. B.1). The joint probability $P(\hat{\dot{p}}_B, \dot{p}_B) = P(\hat{\dot{p}}_B \mid \dot{p}_B)\, P(\dot{p}_B)$, where the conditional probability $P(\hat{\dot{p}}_B \mid \dot{p}_B)$ will be normal with mean $\dot{p}_B$, since p_B is itself an average. The distribution of velocity estimates $\hat{\dot{p}}_B$ is

$$P(\hat{\dot{p}}_B) = \int P(\hat{\dot{p}}_B, \dot{p}_B)\, d\dot{p}_B \tag{B.6}$$

and will always be wider than $P(\dot{p}_B)$ due to convolution with the additional uncertainty $\delta\hat{\dot{p}}_B$. However, when $\delta\hat{\dot{p}}_B \ll \sigma\dot{p}_B$ then this widening can be ignored. Therefore, we need extremely accurate estimates $\hat{p}_B$ in order to calculate $\dot{p}_B$ by finite difference.

We used $N = 1000$ trajectories and $\Delta t = 20$ fs to calculate $\hat{\dot{p}}_B$, for which $\delta\hat{\dot{p}}_B = 0.00112$ / fs. By contrast, the $\hat{\dot{p}}_B$ sampled from $n = 100$ trajectory pairs used to calculate $\kappa[p_B]$ were distributed with $\sigma\dot{p}_B \approx 0.00345$ / fs. The joint probability of the true velocity $\dot{p}_B$ and the velocity estimate $\hat{\dot{p}}_B$ is shown in Fig. B.1(a). Fig. B.1(b) shows that for the calculated $\delta\hat{\dot{p}}_B$ and $\sigma\dot{p}_B$, the distributions $P(\dot{p}_B)$ and $P(\hat{\dot{p}}_B)$ show little difference. Therefore, errors in $\kappa[p_B]$ due to estimation error $\delta\hat{\dot{p}}_B$ are negligible.

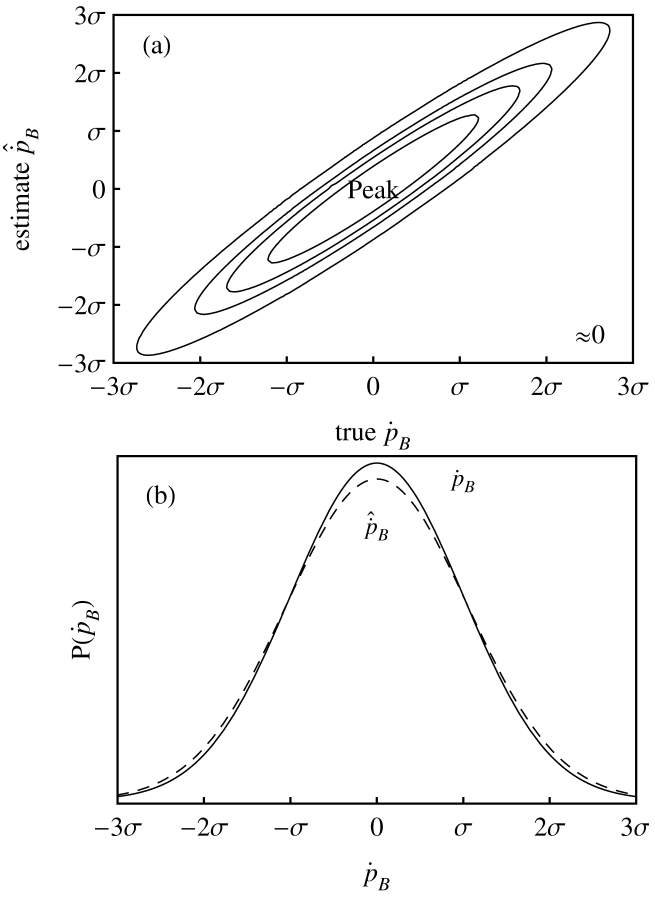

Figure B.1. (a) Joint probability of the true velocity $\dot{p}_B$ and the velocity estimate $\hat{\dot{p}}_B$. The uncertainty in the estimate is independent of the true velocity. **(b)** Probability of the true velocity $\dot{p}_B$ (solid) and the velocity estimate $\hat{\dot{p}}_B$ (dashed). For both (a) and (b) the axes are marked in increments of $\sigma\dot{p}_B = 0.00345$ / fs.

References

1. Chandler, D. "Statistical mechanics of isomerization dynamics in liquids and the transition state approximation." *J. Chem. Phys.* **1978**, *68*, 2959-2970.

Appendix C

List of Trial Coordinates for Ion-pair Dissociation

The following 71 coordinates were calculated for each shooting point and linear combinations of up to 3 parameters were tested using Likelihood Maximization. The number in bold indicates how many versions of the coordinate were tested depending on the parameters described.

- **Ion-pair distance, r_{ion}, 1** – The distance between the sodium cation and chlorine anion.
- **Ion coordination numbers, N_{is}, 4** – The number of water molecules within the *s*th radial shell (2 options) around each ion *i* (2 options) was counted

$$N_{is} = \sum_{w} f_{iws} \tag{C.1}$$

where the sum is over all water molecules. The continuous indicator function

$$f_{iws} = \frac{1 - \tanh\left[a(r_{iw} - b_{is})\right]}{2} \tag{C.2}$$

gives each water a weight between 0 and 1. The subscripts *i*, *w* and *s* refer to an ion, a water molecule, and solvation shell, respectively. The distance r_{iw} is to the oxygen atom when calculating Na^+ coordination and to the closest hydrogen atom when

calculating Cl^- coordination. The parameter b_{is} was set to the s^{th} local minima in the radial distribution function $g(r_{iw})$. For shell s = 1, b_{is} = 3.25 and 3.1 Å for Na^+ and Cl^-, respectively. For shell s = 2, b_{is} = 5.75 Å for both ions. The parameter a = 3. In the manuscript, only shell s = 1 coordination numbers are reported. See Ref. [1].

- **Ion-pair coordination number, N_{ip}, 2** – The number of waters coordinated to the ion pair is

$$N_{ip} = \sum_{w} \max\left(f_{Na,w}, f_{Cl,w}\right) \tag{C.3}$$

N_{ip} is defined for both shell s = 1 and s = 2. See Refs. [2] and [3].

- **Number of bridging waters, N_B, 1** - The number of waters jointly coordinated to both ions simultaneously is

$$N_B = \sum_{w} \min\left(f_{Na,w}, f_{Cl,w}\right) = N_{Na} + N_{Cl} - N_{ip} \tag{C.4}$$

N_B is only defined for shell s = 1. See Ref. [4].

- **Interionic water density, ρ_{ii}, 10** – The number density of water molecules between the two ions was calculated by summing over a Gaussian indicator function

$$\rho_{ii} = \frac{1}{(2\pi\sigma^2)^{3/2}} \sum_{w} \exp\left(-\frac{(\boldsymbol{r}_w - \boldsymbol{r}_{mid})^2}{2\sigma^2}\right) \tag{C.5}$$

where $\mathbf{r}_{mid}$ is the midpoint between the ions, $\mathbf{r}_w$ is the position of the w^{th} water molecule measured by the molecule's a) centroid or b) the center of mass. ρ_{ii} was calculated for the following values of σ: 3.7 Å, $r_{ion}/4$, $r_{ion}/3$, $r_{ion}/2$ and r_{ion} (5 options). Note that except for σ = 3.7 Å, the resolution of the indicator function is always commensurate with the ion-pair distance. The sum in Eq. (C.5) gives the number of

water molecules in a volume $V_{ii} = (2\pi\sigma^2)^{3/2}$. As r_{ion} increases, ρ_{ii} approaches the bulk water density. See Ref [5].

- **Optimization coordinates, 15** – Additional coordinates were calculated after energy minimization with all atoms fixed except for a) the ions, b) the 1st shell water orientations or c) both the ions and 1st shell water orientations. The following coordinates were calculated for each reference configuration: ion-pair distance, the potential energy, the change in ion-pair distance, the change in the potential energy, and the ratio of the potential energy change to the ion-pair distance change. This family of coordinates was inspired by the dynamical caging regime [6].
- **Coulombic potential energy, 4** – We tested the Coulomb interaction between a) all atoms (i.e., the total Coulombic potential energy), b) the water molecules alone, c) the ions and the oxygen atoms on all of the water molecules, d) the ions and the hydrogen atoms on all of the water molecules. The direct ionic Coulombic energy is an identical coordinate to the ion-pair distance. The Coulomb interaction between the ions and all water molecules is an identical coordinate to the energy gap.
- **Energy gap, 3** – The energy gap of Warshel [7] is the energy difference between the shooting point configuration and a reference configuration with the ion charges swapped. The energy gap was calculated using a) all water molecules, b) only waters in the 1st shell, and c) only waters in the 2nd shell. For ion-pair dissociation, the energy gap is an identical coordinate to the interaction Coulombic energy between the ions and water molecules.
- **Tetrahedral water structure coordinates, 16** – The coordinate

$$q_j = \sum_{i=1}^{4} \sum_{k \neq i} \left(\cos(\theta_{ijk}) + \frac{1}{3} \right)^2 \quad \text{(C.6)}$$

was devised by Chau & Hardwick [8] to give $q_j = 0$ if each combination of the 4 nearest neighbors i and k form tetrahedral angles of $\theta_{ijk} = 109.47°$ with the central molecule j. The sum of q_j was tested with the following variations: water molecules j located in the 1st or 2nd shell of either ion (4 options), nearest neighbor location measured at the oxygen or hydrogen position (2 options), restricting the 4 nearest neighbors to water molecules or allowing ions to be one of the neighbors (2 options).

- **Angle: Na-water-Cl, 3** – The Na-water-Cl angles were calculated for every water molecule and we tested a) the largest angle, b) the 2nd largest angle and c) the average of the two largest angles. An angle of 180° corresponds to a water molecule between the ions on the interionic axis. Smaller angles correspond to water molecules further away. See Ref. [9].
- **Angle: Na-Cl-Water, 3** – Cosines of the Na-Cl-water angles were calculated for water molecules in the 1st or 2nd shell around the center atom. For a 1st shell water, cos(0°) = 1 if the water is directly between the ions. We tested a) the maximum cosine (minimum angle) for waters in shell s = 1, b) the maximum cosine for waters in shell s= 2 and c) the Kullback-Leibler divergence between the shell s = 1 and s = 2 distributions. See Ref. [10].
- **Angle: Cl-Na-water, 3** – Same as "Angle: Na-Cl-water" but with sodium the center atom.
- **Hydrogen bonding, 4** – The number of hydrogen bonds between the 1st and 2nd shell waters for each x = Na, Cl, ip, B (see "Coordination numbers").

- **Force on r_{ion}, 2** – We explicitly tested a) the total force exerted on the ions and b) the force exerted on the ions only by water molecules. In addition, we tested the accelerations of all 70 proposed coordinates. These coordinates were inspired by the nonadiabatic regime of Grote-Hynes theory [6].

E. References

1. Ensing, B.; Laio, A.; Gervasio, F. L.; Parrinello, M.; Klein, M. L. "A minimum free energy reaction path for the E2 reaction between fluoro ethane and a fluoride ion." *J. Am. Chem. Soc.* **2004**, *126*, 9492-9493.
2. Fennell, C. J.; Bizjak, A.; Vlachy, V.; Dill, K. A. "Ion Pairing in Molecular Simulations of Aqueous Alkali Halide Solutions." *J. Phys. Chem. B* **2009**, *113*, 6782-6791.
3. Ballard, A. J.; Dellago, C. "Toward the Mechanism of Ionic Dissociation in Water." *J. Phys. Chem. B* **2012**, *116*, 13490-13497.
4. Dang, L. X.; Smith, D. E. "Molecular-Dynamics Simulations of Aqueous Ionic Clusters Using Polarizable Water." *J. Chem. Phys.* **1993**, *99*, 6950-6956.
5. Ciccotti, G.; Ferrario, M.; Hynes, J. T.; Kapral, R. "Dynamics of Ion-Pair Interconversion in a Polar-Solvent." *J. Chem. Phys.* **1990**, *93*, 7137-7147.
6. van der Zwan, G.; Hynes, J. T. "Dynamical Polar-Solvent Effects on Solution Reactions - A Simple Continuum Model." *J. Chem. Phys.* **1982**, *76*, 2993-3001.
7. Warshel, A. "Dynamics of Reactions in Polar-Solvents - Semi-classical Trajectory Studies of Electron-Transfer and Proton-Transfer Reactions." *J. Phys. Chem.* **1982**, *86*, 2218-2224.
8. Chau, P. L.; Hardwick, A. J. "A new order parameter for tetrahedral configurations." *Mol. Phys.* **1998**, *93*, 511-518.
9. DeMille, R. C.; Molinero, V. "Coarse-grained ions without charges: Reproducing the solvation structure of NaCl in water using short-ranged potentials." *J. Chem. Phys.* **2009**, *131*, 034107.
10. Geissler, P. L.; Dellago, C.; Chandler, D. "Kinetic pathways of ion pair dissociation in water." *J. Phys. Chem. B* **1999**, *103*, 3706-3710.

Appendix D

Regimes of Grote-Hynes Theory

In Grote-Hynes theory, the transmission coefficient κ is given by

$$\kappa = \frac{\lambda_r}{\omega_b} \tag{D.1}$$

where ω_b is the frequency of the solvent-equilibrated free energy barrier and the reactive frequency λ_r is the largest, positive root to the Grote-Hynes equation

$$\lambda^2 + \lambda\tilde{\eta}(\lambda) - \omega_b^2 = 0 \tag{D.2}$$

Since $0 \leq \kappa \leq 1$, from Eq. (D.1), $0 \leq \lambda_r \leq \omega_b$, as can also be seen in Fig. D.1. The frequency-dependent friction $\tilde{\eta}(\lambda)$ is the Laplace transform of the time-dependent friction kernel $\eta(t)$,

$$\tilde{\eta}(\lambda) = \int_0^\infty dt\, e^{-\lambda t}\, \eta(t) \tag{D.3}$$

From the fluctuation-dissipation theorem, the time-dependent friction is proportional to the time correlation function of the fluctuating force R in the generalized Langevin equation,

$$\eta(t) = \frac{\langle R(0)R(t)\rangle}{m\, k_B T} \tag{D.4}$$

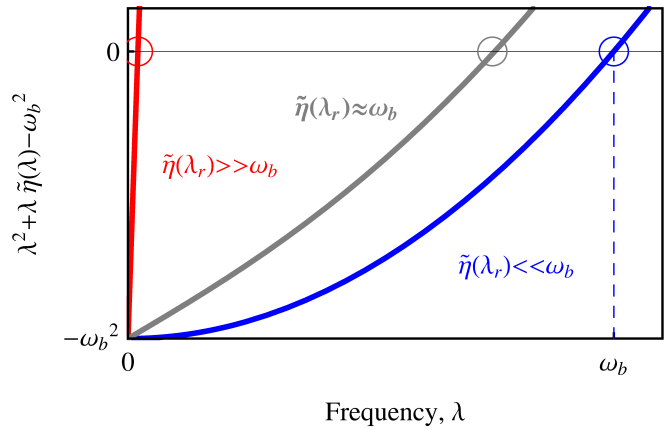

Figure D.1. Schematic showing solutions to the Grote-Hynes equation for strong friction (red), weak friction (blue) and intermediate friction (gray). The reactive frequency λ_r is circled in each case.

where m is a generalized mass associated with the reaction coordinate [1]. Although it is neither strictly monotonic nor positive, the time-dependent friction is typically a decaying function. The frequency-dependent friction thus also decreases with increasing frequency. Typically, the strongest possible friction for the reactive system is the zero-frequency friction $\tilde{\eta}(0)$. Because $\lambda_r \leq \omega_b$, the weakest possible friction is $\tilde{\eta}(\omega_b)$.

The importance of non-Markov effects can be anticipated from the frequency-dependent friction $\tilde{\eta}(\lambda)$ on the range $0 < \lambda < \omega_b$ or equivalently by a comparison of timescales. Markovian friction is frequency independent, $\tilde{\eta}(\omega_b) \approx \tilde{\eta}(0)$, and the transmission coefficient of Kramers theory is recovered. If $\tilde{\eta}(\omega_b) \ll \tilde{\eta}(0)$, then the friction is non-Markovian and the slowest solvent modes are not able to equilibrate on the barrier crossing timescale. $\tilde{\eta}(\omega_b)$ suggests intrinsic timescales for comparison: $\exp(-\omega_b t)$ decays on a timescale ω_b^{-1} and $\eta(t)$ decays on a correlation timescale τ_c [2]

$$\tau_c = \int_0^\infty dt \frac{n(t)}{\eta(0)} = \frac{\tilde{\eta}(0)}{\eta(0)} \tag{D.5}$$

Fig. D.2 compares $\eta(t)$ to $\exp(-\omega_b t)$ for broad and narrow barriers. When the barrier is broad or, equivalently, the solvent response times are fast, $\tau_c \ll \omega_b^{-1}$, then $\exp(-\omega_b t) \approx 1$ over the range of non-zero friction and $\tilde{\eta}(\omega_b) \approx \tilde{\eta}(0)$. When the barrier is narrow or the solvent responds slowly, $\tau_c \gg \omega_b^{-1}$, then $\eta(t) \approx \eta(0)$ over the range that $\exp(-\omega_b t)$ decays and $\tilde{\eta}(\omega_b) \ll \tilde{\eta}(0)$. Therefore, the fast and slow bath regimes can be understood using $\tilde{\eta}(\omega_b)$ and $\tilde{\eta}(0)$ or τ_c and ω_b^{-1}.

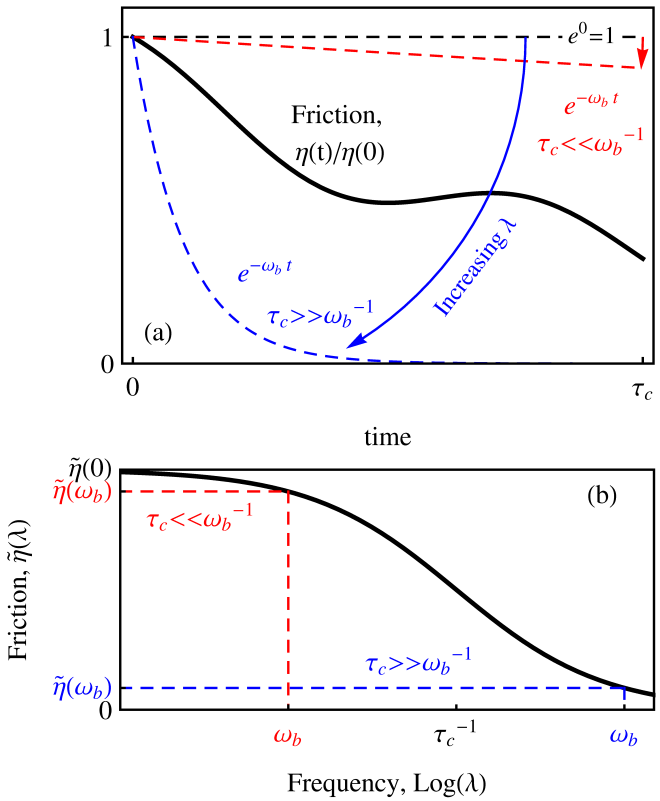

Figure D.2. Schematics showing the comparison of timescales τ_c and ω_b emerges from evaluating the lower limit of the frequency-dependent friction, $\tilde{\eta}(\omega_b)$. **(a)** The decrease in friction (solid) is compared to the decrease in Laplace kernel $\exp(-\lambda t)$ (dashed) at the limits $\lambda = 0$ and $\lambda = \omega_b$. For a broad barrier or fast solvent (red) $\exp(-\omega_b t)$ is nearly constant. For a narrow barrier or slow solvent (blue), $\exp(-\omega_b t)$ decays rapidly while the friction is nearly constant. The solid arrows indicate the range of $\exp(-\lambda t)$ for increasing λ. **(b)** The frequency-dependent friction is nearly constant for the broad barrier (red) showing that the Markovian regime applies. For the narrow barrier (blue), the non-Markovian regime applies because the friction decays on the timescale of barrier crossing.

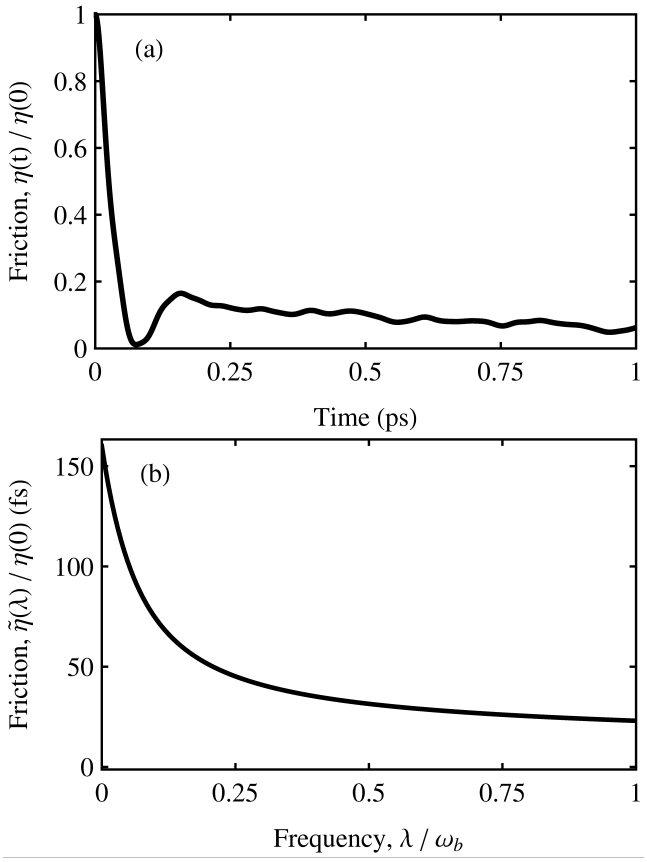

Figure D.3. (a) Time-dependent friction η(t) on r_{ion} normalized by the initial friction η(0) for the ion-pair dissociation model considered in Chapter 4 calculated according to Ref. [3]. **(b)** Frequency-dependent friction $\tilde{\eta}(\lambda)$ divided by the initial time-dependent friction η(0), corresponding to the Laplace transform of (a). The barrier frequency $\omega_b = (60\ \text{fs})^{-1}$. A cutoff of 2 ps was used in integrating η(t), giving $\tau_c = 160$ fs and a frictional decay $\tilde{\eta}(\omega_b) / \tilde{\eta}(0) = 0.144$. The numerical precision of τ_c and $\tilde{\eta}(0)$ from GH theory have no bearing on our conclusions. Rather, we primarily used GH theory to identify which dynamical regimes could be important and to design candidate reaction coordinates accordingly.

The equations of motion effectively change to reflect the different time and friction scales [4]. For example, for the slow bath the equation of motion for an average trajectory becomes

$$\begin{aligned} \ddot{q} &= \omega_b^2 q - \int_0^t d\tau\, \eta(t-\tau)\, \dot{q}(\tau) \\ &= \omega_b^2 q - \eta(0) \int_0^t d\tau\, \dot{q}(\tau) \\ &= \omega_b^2 q - \eta(0) q \end{aligned} \tag{D.6}$$

The two non-Markovian regimes discussed in Chapter 4, nonadiabatic and dynamic caging, are differentiated by the strength of the zero-time or initial friction $\eta(0)$. When the initial friction is weak, $\eta(0) \ll \omega_b^2$, reaction coordinate motion is impeded but not restricted in range. Barrier crossings are completed nonadiabatically, with the solvent modes frozen. The effective barrier frequency $[\omega_b^2 - \eta(0)]^{1/2}$ is then modified from the frictionless barrier frequency. When the initial friction is strong, $\eta(0) \gg \omega_b^2$, Eq. (D.6) describes harmonic motion in an effective potential that restricts the reaction coordinate to values at the barrier top. The location of the effective potential, or cage, moves on the much longer timescale τ_c and so the moving cage will slowly evolve to either product or reactant states.

To consider the effect on κ of increasing non-Markovian friction one must specify how the friction is increased [5]. From Eq. (D.5), the static friction $\tilde{\eta}(0)$ is the product of the initial friction $\eta(0)$ and the correlation time τ_c. This suggests two standard methods for increasing $\tilde{\eta}(0)$: (a) increasing $\eta(0)$ at constant τ_c, or (b) increasing τ_c at constant $\eta(0)$. The Grote-Hynes theory predictions for an exponential friction are shown in Fig. D.4. If friction is increased by method (a) corresponding to horizontal motion in Fig. D.4, the coordinate motion will change from nonadiabatic to dynamically caged and κ will approach zero, $\kappa \sim \tilde{\eta}(0)^{-1}$. If friction is

increased by method (b), corresponding to diagonal motion in Fig. D.4, there is no crossover of these regimes and Grote-Hynes theory predicts κ will plateau.

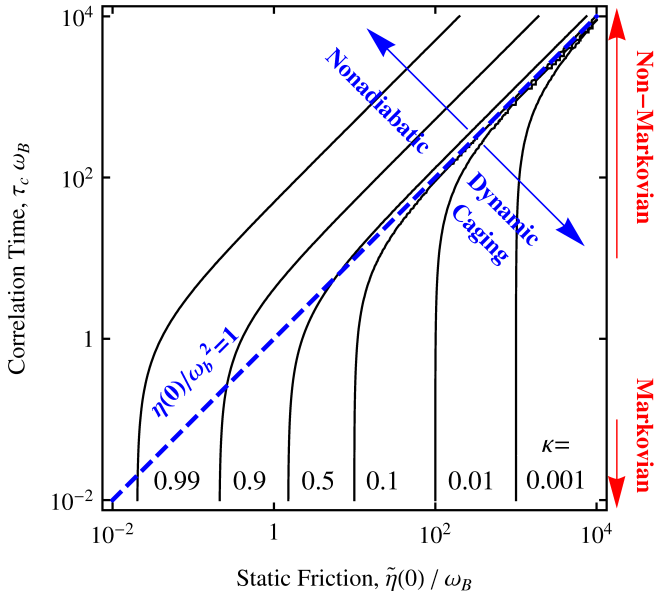

Figure D.4. Contour lines (gray) of the Grote-Hynes transmission coefficient κ for the non-Markovian exponential friction $\eta(t) = \eta(0)\ \exp[-t\ /\ \tau_c]$. The static friction is defined by Eq. (D.5). The nonadiabatic and dynamic caging regimes are separated by the ω_b^2-isosurface of the initial friction (blue, dashed line).

Straub, Berne, and Borkovec [5] investigated reaction rates using direct numerical simulations of the GLE for exponential friction in a double-well potential. Since the reactant and product states were wells, they anticipated—and observed—an energy diffusion limited [6] behavior for weak friction that would cause κ to deviate from Grote-Hynes theory predictions. Note that Grote-Hynes theory was derived for an infinite parabolic barrier potential on which the energy diffusion limit cannot occur. Straub et al. also increased τ_c at constant $\eta(0)$ and found that κ passes through a maximum and then decreases instead of plateauing as Grote-Hynes theory predicts. Similar to the weak friction limit, Straub *et al.* found the disagreement from Grote-Hynes theory at large τ_c was also due to slow energy diffusion, though in an effective potential that evolves with time. Importantly, they reported

this result only for the case of moderate initial friction, $\eta(0) / \omega_b^2$ less than or approximately 1, when the effective potential is a shallow double-well. In the dynamic caging limit, $\eta(0) / \omega_b^2 \gg 1$, the effective potential is a single harmonic well.

References

1. Johannesson, G. H.; Jonsson, H. "Optimization of hyperplanar transition states." *J. Chem. Phys.* **2001**, *115*, 9644-9656.
2. Gertner, B. J.; Wilson, K. R.; Hynes, J. T. "Nonequilibrium Solvation Effects on Reaction-Rates for Model SN2 Reactions in Water." *J. Chem. Phys.* **1989**, *90*, 3537-3558.
3. Rey, R.; Guardia, E. "Dynamic Aspects of the Na+-Cl- Ion-Pair Association in Water." *J. Phys. Chem.* **1992**, *96*, 4712-4718.
4. Kohen, D.; Tannor, D. J. "Phase-Space Distribution Function Formulation of the Method of Reactive Flux - Memory Friction." *J. Chem. Phys.* **1995**, *103*, 6013-6020.
5. Straub, J. E.; Borkovec, M.; Berne, B. J. "Non-Markovian Activated Rate-Processess - Comparison of Current Theories with Numerical-Simulation Data." *J. Chem. Phys.* **1986**, *84*, 1788-1794.
6. Kramers, H. A. "Brownian motion in field of force and diffusion model of chemical reactions." *Physica* **1940**, *7*, 284-304.